PISA 2018 Results

WHAT STUDENTS KNOW AND CAN DO

VOLUME I

Programme for International Student Assessment

OECD

PISA 2018 Results (Volume I)

WHAT STUDENTS KNOW AND CAN DO

OECD
BETTER POLICIES FOR BETTER LIVES

Please cite this publication as:
OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, *https://doi.org/10.1787/5f07c754-en*.

ISBN 978-92-64-46038-6 (print)
ISBN 978-92-64-54188-7 (pdf)
ISBN 978-92-64-45898-7 (HTML)
ISBN 978-92-64-70172-4 (epub)

PISA
ISSN 1990-8539 (print)
ISSN 1996-3777 (online)

Photo credits: Cover
© LuminaStock/iStock
© Dean Mitchell/iStock
© bo1982/iStock
© karandaev/iStock
© IA98/Shutterstock
© Tupungato/Shutterstock

Corrigenda to publications may be found on line at: *www.oecd.org/about/publishing/corrigenda.htm*.

Preface

Among its many findings, our PISA 2018 assessment shows that 15-year-old students in the four provinces/municipalities of China that participated in the study – Beijing, Shanghai, Jiangsu and Zhejiang – outperformed by a large margin their peers from all of the other 78 participating education systems, in mathematics and science. Moreover, the 10% most disadvantaged students in these four jurisdictions also showed better reading skills than those of the average student in OECD countries, as well as skills similar to the 10% most advantaged students in some of these countries. True, these four provinces/municipalities in eastern China are far from representing China as a whole, but the size of each of them compares to that of a typical OECD country, and their combined populations amount to over 180 million. What makes their achievement even more remarkable is that the level of income of these four Chinese regions is well below the OECD average. The quality of their schools today will feed into the strength of their economies tomorrow.

In this context, and given the fact that expenditure per primary and secondary student rose by more than 15% across OECD countries over the past decade, it is disappointing that most OECD countries saw virtually no improvement in the performance of their students since PISA was first conducted in 2000. In fact, only seven of the 79 education systems analysed saw significant improvements in the reading, mathematics and science performance of their students throughout their participation in PISA, and only one of these, Portugal, is a member of the OECD.

During the same period, the demands placed on the reading skills of 15-year-olds have fundamentally changed. The smartphone has transformed the ways in which people read and exchange information; and digitalisation has resulted in the emergence of new forms of text, ranging from the concise, to the lengthy and unwieldy. In the past, students could find clear and singular answers to their questions in carefully curated and government-approved textbooks, and they could trust those answers to be true. Today, they will find hundreds of thousands of answers to their questions on line, and it is up to them to figure out what is true and what is false, what is right and what is wrong. Reading is no longer mainly about extracting information; it is about constructing knowledge, thinking critically and making well-founded judgements. Against this backdrop, the findings from this latest PISA round show that fewer than 1 in 10 students in OECD countries was able to distinguish between fact and opinion, based on implicit cues pertaining to the content or source of the information. In fact, only in the four provinces/municipalities of China, as well as in Canada, Estonia, Finland, Singapore and the United States, did more than one in seven students demonstrate this level of reading proficiency.

There is another side to this. The kinds of things that are easy to teach are nowadays also easy to digitise and automate. In the age of artificial intelligence (AI) we need to think harder about how to develop first-class humans, and how we can pair the AI of computers with the cognitive, social and emotional skills, and values of people. AI will amplify good ideas and good practice in the same way as it amplifies bad ideas and bad practice – it is ethically neutral. However, AI is always in the hands of people who are not neutral. That is why education in the future is not just about teaching people, but also about helping them develop a reliable compass to navigate an increasingly complex, ambiguous and volatile world. Whether AI will destroy or create more jobs will very much depend on whether our imagination, our awareness, and our sense of responsibility will help us harness technology to shape the world for the better. These are issues that the OECD is currently exploring with our Education 2030 project.

PISA is also broadening the range of outcomes that it measures, including global competency in 2018, creative thinking in 2021, and learning in the digital world in 2024. The 2018 assessment asked students to express how they relate to others, what they think of their lives and their future, and whether they believe they have the capacity to grow and improve.

Measuring the well-being of 15-year-old students, the target PISA population, is particularly important, as students at this age are in a key transition phase of physical and emotional development. When it comes to those social and emotional outcomes, the top-performing Chinese provinces/municipalities are among the education systems with most room for improvement.

Even across OECD countries, just about two in three students reported that they are satisfied with their lives, and that percentage shrank by five percentage points between 2015 and 2018. Some 6% of students reported always feeling sad. In almost every education system, girls expressed greater fear of failure than boys, even when they outperformed boys in reading by a large margin. Almost a quarter of students reported being bullied at least a few times a month. Perhaps most disturbingly, in one-third of countries and economies that participated in PISA 2018, including OECD countries such as Greece, Mexico and Poland, more than one in two students said that intelligence was something about them that they couldn't change very much. Those students are unlikely to make the investments in themselves that are necessary to succeed in school and in life. Importantly, having a growth mindset seems consistently associated with students' motivation to master tasks, general self-efficacy, setting learning goals and perceiving the value of school, and negatively associated with their fear of failure. Even if the well-being indicators examined by PISA do not refer specifically to the school context, students who sat the 2018 PISA test cited three main aspects of their lives that influence how they feel: life at school, their relationships with their parents, and how satisfied they are with the way they look.

It may be tempting to conclude that performing better in school will necessarily increase anxiety about schoolwork and undermine students' well-being. But countries such as Belgium, Estonia, Finland and Germany show that high performance and a strong sense of well-being can be achieved simultaneously; they set important examples for others.

Other countries show that equity and excellence can also be jointly achieved. In Australia, Canada, Denmark, Estonia, Finland, Hong Kong (China), Japan, Korea, Macao (China), Norway and the United Kingdom, for example, average performance was higher than the OECD average while the relationship between socio-economic status and reading performance was weaker than the OECD average. Moreover, one in ten disadvantaged students was able to score in the top quarter of reading performance in their country/economy, indicating that poverty is not destiny. The data also show that the world is no longer divided between rich and well-educated nations and poor and badly educated ones. The level of economic development explains just 28% of the variation in learning outcomes across countries if a linear relationship is assumed between the two.

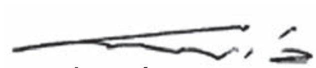
However, it remains necessary for many countries to promote equity with much greater urgency. While students from well-off families will often find a path to success in life, those from disadvantaged families have generally only one single chance in life, and that is a great teacher and a good school. If they miss that boat, subsequent education opportunities will tend to reinforce, rather than mitigate, initial differences in learning outcomes. Against this background, it is disappointing that in many countries a student's or school's post code remains the strongest predictor of their achievement. In Argentina, Bulgaria, the Czech Republic, Hungary, Peru, the Slovak Republic and the United Arab Emirates, a typical disadvantaged student has less than a one-in-eight chance of attending the same school as high achievers.

Furthermore, in over half of the PISA-participating countries and economies, principals of disadvantaged schools were significantly more likely than those of advantaged schools to report that their school's capacity to provide instruction is hindered by a lack or inadequacy of educational material; and in 31 countries and economies, principals of disadvantaged schools were more likely than those of advantaged ones to report that a lack of teaching staff hinders instruction. In these systems, students face a double disadvantage: one that comes from their home background and another that is created by the school system. There can be numerous reasons why some students perform better than others, but those performance differences should never be related to the social background of students and schools.

Clearly, all countries have excellent students, but too few countries have enabled all of their students to excel and fulfill their potential to do so. Achieving greater equity in education is not only a social justice imperative, it is also a way to use resources more effectively, increase the supply of skills that fuel economic growth, and promote social cohesion. For those with the right knowledge and skills, digitalisation and globalisation have been liberating and exciting; for those who are insufficiently prepared, these trends can mean vulnerable and insecure work, and a life with few prospects. Our economies are linked together by global chains of information and goods, but they are also increasingly concentrated in hubs where comparative advantage can be built and renewed. This makes the distribution of knowledge and wealth crucial, and it can only be possible through the distribution of education opportunities.

Equipping citizens with the knowledge and skills necessary to achieve their full potential, to contribute to an increasingly interconnected world, and to convert better skills into better lives needs to become a more central preoccupation of policy makers around the world. Fairness, integrity and inclusiveness in public policy thus all hinge on the skills of citizens. In working to achieve these goals, more and more countries are looking beyond their own borders for evidence of the most successful and efficient education policies and practices.

PISA is not only the world's most comprehensive and reliable indicator of students' capabilities, it is also a powerful tool that countries and economies can use to fine-tune their education policies. Volume V of PISA 2018 Results, which will be published in June 2020, will highlight some of the policies and practices that predict the success of students, schools and education systems. That is why the OECD produces this triennial report on the state of education around the globe: to share evidence of the best policies and practices, and to offer our timely and targeted support to help countries provide the best education possible for all of their students.



Angel Gurría
OECD Secretary-General

Foreword

Up to the end of the 1990s, OECD comparisons of education outcomes were mainly based on measures of years of schooling, which are not reliable indicators of what people are actually able to do. With the Programme for International Student Assessment, PISA, we tried to change this. The transformational idea behind PISA lay in testing the skills of students directly, through a metric that was internationally agreed upon; linking that with data from students, teachers, schools and systems to understand performance differences; and then harnessing the power of collaboration to act on the data, both by creating shared points of reference and by leveraging peer pressure.

The aim with PISA was not to create another layer of top-down accountability, but to help schools and policy makers shift from looking upwards within the bureaucracy towards looking outwards to the next teacher, the next school, the next country. In essence, PISA counts what counts, and makes that information available to educators and policy makers so they can make more informed decisions.

The OECD countries that initiated PISA tried to make PISA different from traditional assessments in other ways too. In a world that rewards individuals increasingly not just for what they know, but for what they can do with what they know, PISA goes beyond assessing whether students can reproduce what they have learned in school. To do well in PISA, students have to be able to extrapolate from what they know, think across the boundaries of subject-matter disciplines, apply their knowledge creatively in novel situations and demonstrate effective learning strategies. If all we do is teach our children what we know, they might remember enough to follow in our footsteps; but if we teach them how to learn, they can go anywhere they want.

Some people argued that the PISA tests are unfair, because they confront students with problems they have not encountered in school. But life is unfair, because the real test in life is not whether we can remember what we learned at school yesterday, but whether we will be able to solve problems that we can't possibly anticipate today.

But the greatest strength of PISA lies in its working methods. Most assessments are centrally planned and then contracted to engineers who build them. That's how tests are created that are owned by a company – but not by the people who are needed to change education. PISA turned that on its head. The idea of PISA attracted the world's best thinkers and mobilised hundreds of experts, educators and scientists from the participating countries to build a global assessment. Today, we would call that crowdsourcing; but whatever we call it, it created the ownership that was critical for success.

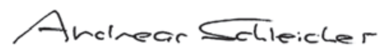
In a nutshell, PISA owes its success to a collaborative effort between the participating countries and economies, the national and international experts and institutions working within the framework of the PISA Consortium, and the OECD Secretariat. Countless subject-matter experts, practitioners and policy makers from the participating countries worked tirelessly to build agreement on which learning outcomes are important to measure and how to measure them best; to design and validate assessment tasks that can reflect those measures adequately and accurately across countries and cultures; and to find ways to compare the results meaningfully and reliably. The OECD Secretariat co-ordinated this effort and worked with countries to make sense of the results and compile this report.

Over the past two decades, PISA has become the world's premier yardstick for evaluating the quality, equity and efficiency of school systems, and an influential force for education reform. It has helped policy makers lower the cost of political action by backing difficult decisions with evidence – but it has also raised the political cost of inaction by exposing areas where policy and practice are unsatisfactory. Today, PISA brings together more than 90 countries, representing 80% of the world economy, in a global conversation about education.

While measurement is the means, the purpose of PISA is to help countries look outwards and incorporate the results of that learning into policy and practice. That outward-looking perspective also seems to be a common trait of many high-performing education systems: they are open to the world and ready to learn from and with the world's education leaders; they do not feel threatened by alternative ways of thinking.

In the end, the laws of physics apply. If we stop pedalling, not only will we not move forward, our bicycles will stop moving at all and will fall over – and we will fall with them. Against strong headwinds, we need to push ourselves even harder. But in the face of challenges and opportunities as great as any that have gone before, human beings need not be passive or inert.

We have agency, the ability to anticipate and the power to frame our actions with purpose. The best-performing PISA countries show us that high-quality and equitable education is an attainable goal, that it is within our means to deliver a future for millions of learners who currently do not have one, and that our task is not to make the impossible possible, but to make the possible attainable.



Andreas Schleicher
Director for Education and Skills
Special Advisor on Education Policy
to the Secretary-General

Acknowledgements

This report is the product of a collaborative effort between the countries and economies participating in PISA, the national and international experts and institutions working within the framework of the PISA Consortium, and the OECD Secretariat.

The development of this volume was guided by Andreas Schleicher and Yuri Belfali and managed by Miyako Ikeda. This volume was drafted by Francesco Avvisati with Jeffrey Mo and edited by Marilyn Achiron. Statistical and analytical support was provided by Guillaume Bousquet and Giannina Rech. Nicolás Miranda led the development of system-level indicators. Rebecca Tessier co-ordinated production with Alison Burke's support and Fung Kwan Tam designed the publication. Juliet Evans and Julia Himstedt provided communication support. Administrative support was provided by Thomas Marwood and Hanna Varkki. This volume also benefitted from the input and expertise of many more OECD staff members who worked on PISA 2018 at various stages of the project. Their names are listed in Annex E of this volume. Many reviewers provided feedback on earlier chapter drafts; their help in improving this volume is gratefully acknowledged.

To support the technical implementation of PISA, the OECD contracted an international consortium of institutions and experts, led by Irwin Kirsch at the Educational Testing Service (ETS). Overall co-ordination of the PISA 2018 assessment, the development of instruments, and scaling and analysis were managed by Claudia Tamassia at ETS. The development of the reading and questionniares frameworks was facilitated by Pearson, led by John de Jong, Peter Foltz and Christine Rozunick. Sampling and weighting services were provided by Westat, led by Keith Rust. Linguistic quality control and the development of the French source version were under the responsibility of cApStAn, led by Steve Dept.

Jean-François Rouet chaired the expert group that guided the preparation of the reading assessment framework and instruments. This group included Paul van den Broek, Kevin Kien Hoa Chung, Dominique Lafontaine, John Sabatini, Sascha Schroeder and Sari Sulkunen. Fons J. R. van de Vijver chaired the expert group that guided the preparation of the questionnaire framework and instruments. This group included Dominique Lafontaine, David Kaplan, Sarah Howie, Andrew Elliot and Therese Hopfenbeck. Keith Rust chaired the Technical Advisory Group, whose members include Theo Eggen, John de Jong, Jean Dumais, Cees Glas, David Kaplan, Kit-Tai Hau, Irwin Kirsch, Oliver Lüdtke, Christian Monseur, Sophia Rabe-Hesketh, Thierry Rocher, Leslie A. Rutkowski, Matthias von Davier, Margaret Wu and Kentaro Yamamoto.

The development of the report was steered by the PISA Governing Board, chaired by Michele Bruniges (Australia), with Peggy Carr (United States), Jimin Cho (Korea) and Carmen Tovar Sánchez (Spain) as vice chairs. Annex E of this volume lists the members of the various PISA bodies, including Governing Board members and National Project Managers in participating countries and economies, the PISA Consortium, and the individual experts and consultants who have contributed to PISA 2018.

PISA

Table of contents

BOXES

FIGURES

TABLES

Executive Summary

Reading proficiency is essential for a wide variety of human activities – from following instructions in a manual; to finding out the who, what, when, where and why of an event; to communicating with others for a specific purpose or transaction. PISA recognises that evolving technologies have changed the ways people read and exchange information, whether at home, at school or in the workplace. Digitalisation has resulted in the emergence and availability of new forms of text, ranging from the concise (text messages; annotated search-engine results) to the lengthy (tabbed, multipage websites; newly accessible archival material scanned from microfiches). In response, education systems are increasingly incorporating digital (reading) literacy into their programmes of instruction.

Reading was the main subject assessed in PISA 2018. The PISA 2018 reading assessment, which was delivered on computer in most of the 79 countries and economies that participated, 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 measure trends in reading literacy over the past two decades. PISA 2018 defined reading literacy as understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.

WHAT STUDENTS KNOW AND CAN DO: MAIN FINDINGS

In reading

- Beijing, Shanghai, Jiangsu and Zhejiang (China) and Singapore scored significantly higher in reading than all other countries/economies that participated in PISA 2018. Estonia, Canada, Finland and Ireland were the highest-performing OECD countries in reading.
- 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.

In mathematics and science

- On average across OECD countries, 76% of students attained Level 2 or higher in mathematics. At a minimum, these students can interpret and recognise, without direct instructions, how a (simple) situation can be represented mathematically (e.g. comparing the total distance across two alternative routes, or converting prices into a different currency). However, in 24 countries and economies, more than 50% of students scored below this level of proficiency.
- Around one in six 15-year-old students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (16.5%), and about one in seven students in Singapore (13.8%), scored at Level 6 in mathematics, the highest level of proficiency that PISA describes. These students are capable of advanced mathematical thinking and reasoning. On average across OECD countries, only 2.4% of students scored at this level.
- On average across OECD countries, 78% of students attained Level 2 or higher in science. At a minimum, these students can recognise the correct explanation for familiar scientific phenomena and can use such knowledge to identify, in simple cases, whether a conclusion is valid based on the data provided. More than 90% of students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (97.9%), Macao (China) (94.0%), Estonia (91.2%) and Singapore (91.0%) achieved this benchmark.

Trends in performance

- On average across OECD countries, mean performance in reading, mathematics and science remained stable between 2015 and 2018.

- There were large differences between individual countries and economies in how their performance changed between 2015 and 2018. For example, mean performance in mathematics improved in 13 countries/economies (Albania, Iceland, Jordan, Latvia, Macao [China], Montenegro, Peru, Poland, Qatar, the Republic of North Macedonia, the Slovak Republic, Turkey and the United Kingdom), declined in 3 countries/economies (Malta, Romania and Chinese Taipei), and remained stable in the remaining 47 participating countries/economies.
- Seven countries/economies saw improvements, on average, in the reading, mathematics and science performance of their students throughout their participation in PISA: Albania, Colombia, Macao (China), the Republic of Moldova, Peru, Portugal and Qatar. Seven countries saw declining mean performance across all three subjects: Australia, Finland, Iceland, Korea, the Netherlands, New Zealand and the Slovak Republic.
- Between 2003 and 2018, Brazil, Indonesia, Mexico, Turkey and Uruguay enrolled many more 15-year-olds in secondary education without sacrificing the quality of the education provided.

Around the world, the share of 15-year-old students, in grade 7 and above, who reached a minimum level of proficiency in reading (at least Level 2 on the PISA scale) ranged from close to 90% in Beijing, Shanghai, Jiangsu and Zhejiang (China), Estonia, Macao (China) and Singapore, to less than 10% in Cambodia, Senegal and Zambia (countries that participated in the PISA for Development assessment in 2017). The share of 15-year-old students who attained minimum levels of proficiency in mathematics (at least Level 2) varied even more – between 98% in Beijing, Shanghai, Jiangsu and Zhejiang (China) and 2% in Zambia. On average across OECD countries, around one in four 15-year-olds did not attain a minimum level of proficiency in reading or mathematics. These numbers show that all countries still have some way to go towards reaching the global goals for quality education, as defined in the UN Sustainable Development Goal for education, by 2030.

Table I.1 [1/2] **Snapshot of performance in reading, mathematics and science**

Countries/economies with a mean performance/share of **top performers above** the OECD average
Countries/economies with a share of **low achievers below** the OECD average

Countries/economies with a mean performance/share of top performers/share of low achievers **not significantly different** from the OECD average

Countries/economies with a mean performance/share of **top performers below** the OECD average
Countries/economies with a share of **low achievers above** the OECD average

	Mean score in PISA 2018			Long-term trend: Average rate of change in performance, per three-year-period			Short-term change in performance (PISA 2015 to PISA 2018)			Top-performing and low-achieving students	
	Reading	Mathematics	Science	Reading	Mathematics	Science	Reading	Mathematics	Science	Share of top performers in at least one subject (Level 5 or 6)	Share of low achievers in all three subjects (below Level 2)
	Mean	Mean	Mean	Score dif.	Score dif.	Score dif.	Score dif.	Score dif.	Score dif.	%	%
OECD average	487	489	489	0	-1	-2	-3	2	-2	15.7	13.4
Estonia	523	523	530	**6**	**2**	0	4	4	-4	22.5	4.2
Canada	520	512	518	-2	**-4**	**-3**	-7	-4	**-10**	24.1	6.4
Finland	520	507	522	**-5**	**-9**	**-11**	-6	-4	**-9**	21.0	7.0
Ireland	518	500	496	0	0	**-3**	-3	-4	-6	15.4	7.5
Korea	514	526	519	**-3**	**-4**	**-3**	-3	2	3	26.6	7.5
Poland	512	516	511	**5**	**5**	2	6	**11**	**10**	21.2	6.7
Sweden	506	502	499	**-3**	-2	-1	6	8	6	19.4	10.5
New Zealand	506	494	508	**-4**	**-7**	**-6**	-4	-1	-5	20.2	10.9
United States	505	478	502	0	-1	2	8	9	6	17.1	12.6
United Kingdom	504	502	505	2	1	-2	6	**9**	-5	19.4	9.0
Japan	504	527	529	1	0	-1	**-12**	-5	**-9**	23.3	6.4
Australia	503	491	503	**-4**	**-7**	**-7**	0	-3	**-7**	18.9	11.2
Denmark	501	509	493	1	-1	0	1	-2	**-9**	15.8	8.1
Norway	499	501	490	1	2	1	**-14**	-1	**-8**	17.8	11.3
Germany	498	500	503	**3**	0	**-4**	-11	-6	-6	19.1	12.8
Slovenia	495	509	507	2	2	**-2**	**-10**	-1	**-6**	17.3	8.0
Belgium	493	508	499	-2	**-4**	**-3**	-6	1	-3	19.4	12.5
France	493	495	493	0	**-3**	-1	-7	2	-2	15.9	12.5
Portugal	492	492	492	**4**	**6**	**4**	-6	1	**-9**	15.2	12.6
Czech Republic	490	499	497	0	**-4**	**-4**	3	7	4	16.6	10.5
Netherlands	485	519	503	**-4**	**-4**	**-6**	**-18**	7	-5	21.8	10.8
Austria	484	499	490	-1	-2	**-6**	0	2	-5	15.7	13.5
Switzerland	484	515	495	-1	**-2**	**-4**	-8	-6	**-10**	19.8	10.7
Latvia	479	496	487	2	2	-1	**-9**	**14**	-3	11.3	9.2
Italy	476	487	468	0	**5**	-2	-8	-3	**-13**	12.1	13.8
Hungary	476	481	481	-1	**-3**	**-7**	6	4	4	11.3	15.5
Lithuania	476	481	482	2	-1	**-3**	3	3	7	11.1	13.9
Iceland	474	495	475	**-4**	**-5**	**-5**	-8	**7**	2	13.5	13.7
Israel	470	463	462	**6**	**6**	3	-9	-7	-4	15.2	22.1
Luxembourg	470	483	477	-1	-2	-2	**-11**	-2	**-6**	14.4	17.4
Turkey	466	454	468	2	**4**	**6**	**37**	**33**	**43**	6.6	17.1
Slovak Republic	458	486	464	**-3**	**-4**	**-8**	5	**11**	3	12.8	16.9
Greece	457	451	452	-2	0	**-6**	-10	-2	-3	6.2	19.9
Chile	452	417	444	**7**	1	1	-6	-5	-3	3.5	23.5
Mexico	420	409	419	2	**3**	2	-3	1	3	1.1	35.0
Colombia	412	391	413	**7**	**5**	**6**	**-13**	1	-2	1.5	39.9
Spain	m	481	483	m	0	-1	m	-4	**-10**	m	m

Notes: Values that are statistically significant are marked in bold (see Annex A3).
Long-term trends are reported for the longest available period since PISA 2000 for reading, PISA 2003 for mathematics and PISA 2006 for science.
Results based on reading performance are reported as missing for Spain (see Annex A9). The OECD average does not include Spain in these cases.
Countries and economies are ranked in descending order of the mean reading score in PISA 2018.
Source: OECD, PISA 2018 Database, Tables I.B1.10, I.B1.11, I.B1.12, I.B1.26 and I.B1.27.
StatLink https://doi.org/10.1787/888934028140

...

Table I.1 [2/2] **Snapshot of performance in reading, mathematics and science**

- [blue] Countries/economies with a mean performance/share of **top performers above** the OECD average
  Countries/economies with a share of **low achievers below** the OECD average
- [white] Countries/economies with a mean performance/share of top performers/share of low achievers **not significantly different** from the OECD average
- [grey] Countries/economies with a mean performance/share of **top performers below** the OECD average
  Countries/economies with a share of **low achievers above** the OECD average

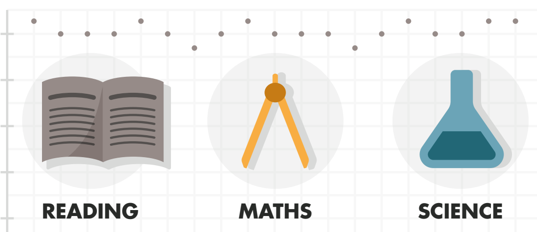
	Mean score in PISA 2018			Long-term trend: Average rate of change in performance, per three-year-period			Short-term change in performance (PISA 2015 to PISA 2018)			Top-performing and low-achieving students	
	Reading	Mathematics	Science	Reading	Mathematics	Science	Reading	Mathematics	Science	Share of top performers in at least one subject (Level 5 or 6)	Share of low achievers in all three subjects (below Level 2)
	Mean	Mean	Mean	Score dif.	Score dif.	Score dif.	Score dif.	Score dif.	Score dif.	%	%
OECD average	487	489	489	0	-1	-2	-3	2	-2	15.7	13.4
Partners											
B-S-J-Z (China)	555	591	590	m	m	m	m	m	m	49.3	1.1
Singapore	549	569	551	**6**	1	**3**	**14**	5	-5	43.3	4.1
Macao (China)	525	558	544	**6**	**6**	**8**	**16**	**14**	**15**	32.8	2.3
Hong Kong (China)	524	551	517	2	0	**-8**	-2	3	-7	32.3	5.3
Chinese Taipei	503	531	516	1	**-4**	-2	6	**-11**	**-17**	26.0	9.0
Croatia	479	464	472	1	0	**-5**	-8	0	-3	8.5	14.1
Russia	479	488	478	**7**	**5**	0	**-16**	-6	**-9**	10.8	11.2
Belarus	474	472	471	m	m	m	m	m	m	9.0	15.9
Ukraine	466	453	469	m	m	m	m	m	m	7.5	17.5
Malta	448	472	457	2	**4**	-1	2	**-7**	**-8**	11.3	22.6
Serbia	439	448	440	**8**	**3**	1	m	m	m	6.7	24.7
United Arab Emirates	432	435	434	-1	4	-2	-2	7	-3	8.3	30.1
Romania	428	430	426	**7**	**5**	2	-6	**-14**	-9	4.1	29.8
Uruguay	427	418	426	1	-2	0	-9	0	**-10**	2.4	31.9
Costa Rica	426	402	416	**-7**	-3	**-6**	-1	2	-4	0.9	33.5
Cyprus	424	451	439	**-12**	**6**	1	**-18**	**14**	**6**	5.9	25.7
Moldova	424	421	428	**14**	**9**	**6**	8	1	0	3.2	30.5
Montenegro	421	430	415	**8**	**8**	2	-6	**12**	4	2.3	31.5
Bulgaria	420	436	424	1	**6**	-1	-12	-5	**-22**	5.5	31.9
Jordan	419	400	429	**4**	3	1	11	**20**	**21**	1.4	28.4
Malaysia	415	440	438	2	**13**	**7**	m	m	m	2.7	27.8
Brazil	413	384	404	3	**5**	2	6	6	3	2.5	43.2
Brunei Darussalam	408	430	431	m	m	m	m	m	m	4.3	37.6
Qatar	407	414	419	**22**	**23**	**18**	5	**12**	2	4.8	37.4
Albania	405	437	417	**10**	**20**	**11**	0	**24**	**-10**	2.5	29.7
Bosnia and Herzegovina	403	406	398	m	m	m	m	m	m	1.0	41.3
Argentina	402	379	404	-1	-1	3	m	m	m	1.2	41.4
Peru	401	400	404	**14**	**12**	**13**	3	**13**	8	1.4	42.8
Saudi Arabia	399	373	386	m	m	m	m	m	m	0.3	45.4
Thailand	393	419	426	**-4**	0	1	**-16**	3	4	2.7	34.6
North Macedonia	393	394	413	1	**23**	**29**	**41**	**23**	**29**	1.7	39.0
Baku (Azerbaijan)	389	420	398	m	m	m	m	m	m	2.1	38.9
Kazakhstan	387	423	397	-1	**5**	-3	m	m	m	2.2	37.7
Georgia	380	398	383	4	**8**	**6**	**-22**	-6	**-28**	1.2	48.7
Panama	377	353	365	2	-2	-4	m	m	m	0.3	59.5
Indonesia	371	379	396	1	2	3	**-26**	-7	-7	0.6	51.7
Morocco	359	368	377	m	m	m	m	m	m	0.1	60.2
Lebanon	353	393	384	m	m	m	7	-3	-3	2.6	49.1
Kosovo	353	366	365	m	m	m	6	4	**-14**	0.1	66.0
Dominican Republic	342	325	336	m	m	m	**-16**	-3	4	0.1	75.5
Philippines	340	353	357	m	m	m	m	m	m	0.2	71.8

Notes: Values that are statistically significant are marked in bold (see Annex A3).
Long-term trends are reported for the longest available period since PISA 2000 for reading, PISA 2003 for mathematics and PISA 2006 for science.
Results based on reading performance are reported as missing for Spain (see Annex A9). The OECD average does not include Spain in these cases.
Countries and economies are ranked in descending order of the mean reading score in PISA 2018.
Source: OECD, PISA 2018 Database, Tables I.B1.10, I.B1.11, I.B1.12, I.B1.26 and I.B1.27.
StatLink https://doi.org/10.1787/888934028140

600 000 students

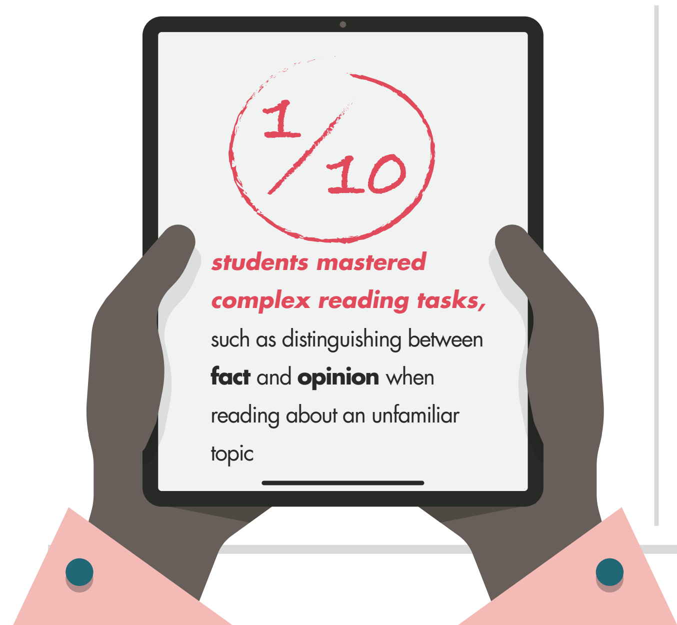
representing about **32 million** 15-year-olds in the schools of the **79 participating countries and economies** sat the **2-hour** PISA test in 2018

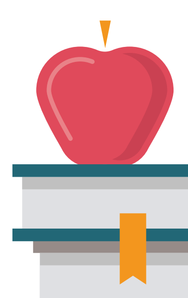
Mean performance in the following subjects did not change over the past 2 decades



But Albania, Estonia, Macao (China), Peru and Poland **saw improvements in at least 2 subjects**

Between 2003 and 2018, Brazil, Indonesia, Mexico, Turkey and Uruguay **enrolled many more 15-year-olds** in secondary education **without sacrificing the quality of the education provided**





1 in 4

students had difficulty with basic aspects of reading, such as identifying the **main idea** in a text of moderate length or **connecting pieces of information** provided by different sources

All data refer to OECD average unless otherwise indicated

PISA

Reader's Guide

Data underlying the figures

The data referred to in this volume are presented in Annex B and, in greater detail, including additional tables, on the PISA website (www.oecd.org/pisa).

Five symbols are used to denote missing data:

a The category does not apply in the country concerned or economy; data are therefore missing.

c There were too few observations to provide reliable estimates (i.e. there were fewer than 30 students or fewer than 5 schools with valid data).

m Data are not available. There was no observation in the sample; these data were not collected by the country or economy; or these data were collected but subsequently removed from the publication for technical reasons.

w Results were withdrawn at the request of the country or economy concerned.

x Data included in another category or column of the table (e.g. x(2) means that data are included in Column 2 of the table).

Coverage

This publication features data on 79 countries and economies, including all OECD Member countries and more than 40 non-OECD Member countries and economies (see map of PISA countries and economies in "What is PISA?").

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.

Notes on Cyprus:

- **Note by Turkey:** The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue".
- **Note by all the European Union Member States of the OECD and the European Union:** The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

B-S-J-Z (China) refers to the four PISA-participating provinces/municipalities of the People's Republic of China (hereafter "China"): Beijing, Shanghai, Jiangsu and Zhejiang.

Data for Viet Nam are included in most tables in Annex B, but not included in tables, figures and texts that report comparisons of performance with other countries and economies' or over time, because full international comparability of results could not be assured at the time this report was published (see Annexes A4 and A6).

International averages

The OECD average corresponds to the arithmetic mean of the respective country estimates. It was calculated for most indicators presented in this report.

The OECD total takes the OECD Member countries as a single entity, to which each country contributes in proportion to the number of 15-year-olds enrolled in its schools. It can be used to assess how an OECD Member country compares with the OECD area as a whole.

On 25 May 2018, the OECD Council invited Colombia to become a Member. While Colombia is included in the OECD averages reported in this publication, at the time of its preparation, Colombia was in the process of completing its domestic procedures for ratification and the deposit of Colombia's instrument of accession to the OECD Convention was pending.

In this publication, the OECD average is generally used when the focus is on comparing performance across education systems. In the case of some countries, data may not be available for specific indicators, or specific categories may not apply. Readers should, therefore, keep in mind that the terms "OECD average" and "OECD total" refer to the OECD Member countries included in the respective comparisons. In cases where data are not available or do not apply for all sub-categories of a given population or indicator, the "OECD average" is not necessarily computed on a consistent set of countries across all columns of a table.

In analyses involving data from multiple years, the OECD average is always reported on consistent sets of OECD Member countries, and several averages may be reported in the same table. For instance, the "OECD average-37" refers to the average across all 36 OECD Member countries (and Colombia), and is reported as missing if fewer than 36 OECD Member countries (and Colombia) have comparable data; the "OECD average-30" includes only 30 OECD Member countries that have non-missing values across all the assessments for which this average itself is non-missing. This restriction allows for valid comparisons of the OECD average over time.

The number in the label used in figures and tables indicates the number of countries included in the average:

- **OECD average-37**: Arithmetic mean across all OECD Member countries (and Colombia).
- **OECD average-36a**: Arithmetic mean across all OECD Member countries (and Colombia), excluding Spain.
- **OECD average-36b**: Arithmetic mean across all OECD Member countries (and Colombia), excluding Austria.
- **OECD average-35a**: Arithmetic mean across all OECD Member countries (and Colombia), excluding Austria and Spain.
- **OECD average-35b**: Arithmetic mean across all OECD Member countries (and Colombia), excluding Spain and the United States.
- **OECD average-30**: Arithmetic mean across all OECD Member countries, excluding Chile, Colombia, Estonia, Israel, Lithuania, Slovenia and the United Kingdom
- **OECD average-29a**: Arithmetic mean across all OECD Member countries, excluding Austria, Chile, Colombia, Estonia, Israel, Lithuania, Slovenia and the United Kingdom
- **OECD average-29b**: Arithmetic mean across all OECD Member countries, excluding Chile, Colombia, Estonia, Israel, Lithuania, Slovenia, Spain and the United Kingdom
- **OECD average-27**: Arithmetic mean across all OECD Member countries, excluding Colombia, Estonia, Lithuania, Luxembourg, the Netherlands, the Slovak Republic, Slovenia, Spain, Turkey and the United Kingdom.
- **OECD average-23**: Arithmetic mean across all OECD Member countries, excluding Austria, Chile, Colombia, Estonia, Israel, Lithuania, Luxembourg, the Netherlands, the Slovak Republic, Slovenia, Spain, Turkey, the United Kingdom and the United States.

Rounding figures

Because of rounding, some figures in tables may not add up exactly to the totals. Totals, differences and averages are always calculated on the basis of exact numbers and are rounded only after calculation.

All standard errors in this publication have been rounded to one or two decimal places. Where the value 0.0 or 0.00 is shown, this does not imply that the standard error is zero, but that it is smaller than 0.05 or 0.005, respectively.

Reporting student data

The report uses "15-year-olds" as shorthand for the PISA target population. PISA covers students who are aged between 15 years 3 months and 16 years 2 months at the time of assessment and who are enrolled in school and have completed at least 6 years of formal schooling, regardless of the type of institution in which they are enrolled, and whether they are in full-time or part-time education, whether they attend academic or vocational programmes, and whether they attend public or private schools or foreign schools within the country.

Reporting school data

The principals of the schools in which students were assessed provided information on their schools' characteristics by completing a school questionnaire. Where responses from school principals are presented in this publication, they are weighted so that they are proportionate to the number of 15-year-olds enrolled in the school.

Focusing on statistically significant differences

This volume discusses only statistically significant differences or changes. These are denoted in darker colours in figures and in bold font in tables. Unless otherwise specified, the significance level is set to 5%. See Annex A3 for further information.

Abbreviations used in this report

ESCS	PISA index of economic, social and cultural status
GDP	Gross domestic product
ICT	Information and communications technology
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
PPP	Purchasing power parity
Score dif.	Score-point difference
S.D.	Standard deviation
S.E.	Standard error
STEM	Science, technology, engineering and mathematics
% dif.	Percentage-point difference

Further documentation

For further information on the PISA assessment instruments and the methods used in PISA, see the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

StatLink

This report has *StatLinks* at the bottom of tables and graphs. To download the matching Excel® spreadsheet, just type the link into your Internet browser, starting with the *https://doi.org* prefix, or click on the link from the e-book version.

Reference

OECD (forthcoming), *PISA 2018 Technical Report,* OECD Publishing, Paris. [1]

PISA

What is PISA?

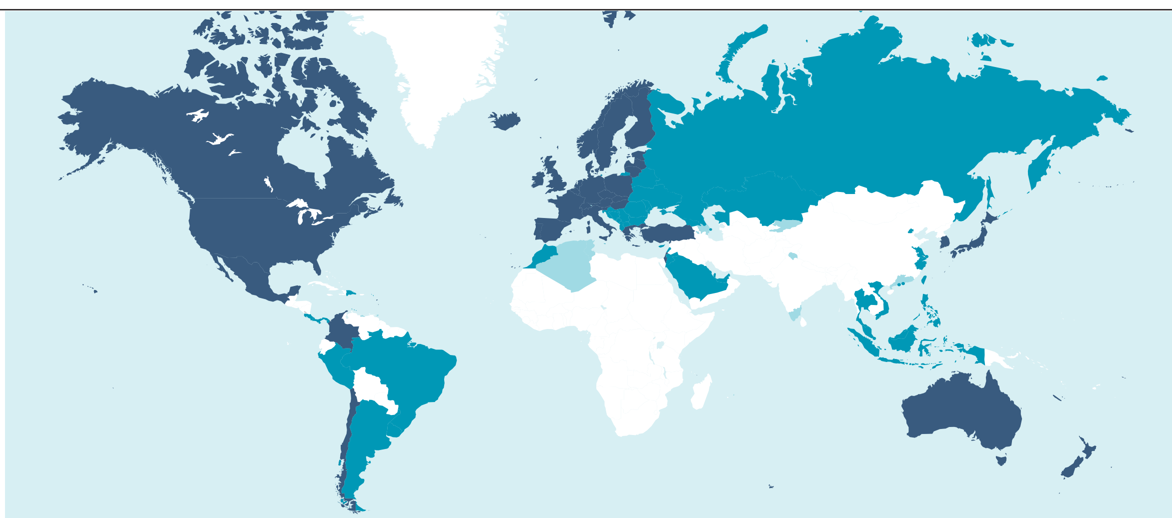
PISA is a triennial survey of 15-year-old students around the world that assesses the extent to which they have acquired key knowledge and skills essential for full participation in social and economic life. PISA assessments do not just ascertain whether students near the end of their compulsory education can reproduce what they have learned; they also examine how well students can extrapolate from what they have learned and apply their knowledge in unfamiliar settings, both in and outside of school.

WHAT IS UNIQUE ABOUT PISA?

PISA is unique because of its:

- **policy orientation**, which links data on student learning outcomes with data on students' backgrounds and attitudes towards learning, and with key factors that shape their learning, in and outside of school; by doing so, PISA can highlight differences in performance and identify the characteristics of students, schools and education systems that perform well
- **innovative concept of "literacy"**, which refers to students' capacity to apply their knowledge and skills in key areas, and to analyse, reason and communicate effectively as they identify, interpret and solve problems in a variety of situations
- **relevance to lifelong learning**, as PISA asks students to report on their motivation to learn, their beliefs about themselves, and their learning strategies
- **regularity**, which enables countries to monitor their progress in meeting key learning objectives
- **breadth of coverage**, which, in PISA 2018, encompassed all 37 OECD countries and 42 partner countries and economies.

Map of PISA countries and economies



OECD member countries

Australia
Austria
Belgium
Canada
Chile
Colombia
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Israel
Italy
Japan
Korea
Latvia
Lithuania
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Slovenia
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States*

Partner countries and economies in PISA 2018

Albania
Argentina
Baku (Azerbaijan)
Belarus
Bosnia and Herzegovina
Brazil
Brunei Darussalam
B-S-J-Z (China)**
Bulgaria
Costa Rica
Croatia
Cyprus
Dominican Republic
Georgia
Hong Kong (China)
Indonesia
Jordan
Kazakhstan
Kosovo
Lebanon
Macao (China)
Malaysia
Malta
Republic of Moldova
Montenegro
Morocco
Republic of North Macedonia
Panama
Peru
Philippines
Qatar
Romania
Russian Federation
Saudi Arabia
Serbia
Singapore
Chinese Taipei
Thailand
Ukraine
United Arab Emirates
Uruguay
Viet Nam

Partner countries and economies in previous cycles

Algeria
Azerbaijan
Guangdong (China)
Himachal Pradesh (India)
Kyrgyzstan
Liechtenstein
Mauritius
Miranda (Venezuela)
Tamil Nadu (India)
Trinidad and Tobago
Tunisia

* Puerto Rico participated in the PISA 2015 assessment (as an unincorporated territory of the United States).

** B-S-J-Z (China) refers to four PISA 2018 participating Chinese provinces/municipalities: Beijing, Shanghai, Jiangsu and Zhejiang. In PISA 2015, the four PISA participating Chinese provinces/municipalities were: Beijing, Shanghai, Jiangsu and Guangdong.

WHICH COUNTRIES AND ECONOMIES PARTICIPATE IN PISA?

PISA is used as an assessment tool in many regions around the world. It was implemented in 43 countries and economies in the first assessment (32 in 2000 and 11 in 2002), 41 in the second assessment (2003), 57 in the third assessment (2006), 75 in the fourth assessment (65 in 2009 and 10 in 2010), 65 in the fifth assessment (2012) and 72 in the sixth assessment (2015). In 2018, 79 countries and economies participated in PISA.

WHAT DOES THE TEST MEASURE?

In each round of PISA, one subject is tested in detail, taking up nearly half of the total testing time. The main subject in 2018 was reading, as it was in 2000 and 2009. Mathematics was the main subject in 2003 and 2012, while science was the main subject in 2006 and 2015. With this alternating schedule, a thorough analysis of achievement in each of the three core subjects is presented every nine years; an analysis of trends is offered every three years.

The *PISA 2018 Assessment and Analytical Framework* (OECD, 2019[11]) presents definitions and more detailed descriptions of the subjects assessed in PISA 2018:

- Reading literacy is defined as students' capacity to understand, use, evaluate, reflect on and engage with texts in order to achieve one's goals, develop one's knowledge and potential, and participate in society.
- Mathematics literacy is defined as students' capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena.
- Science literacy is defined as the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

Box A **Key features of PISA 2018**

The content

- The PISA 2018 survey focused on reading, with mathematics, science and global competence as minor areas of assessment. PISA 2018 also included an assessment of young people's financial literacy, which was optional for countries and economies.

The students

- Some 600 000 students completed the assessment in 2018, representing about 32 million 15-year-olds in the schools of the 79 participating countries and economies.

The assessment

- Computer-based tests were used in most countries, with assessments lasting a total of two hours. In reading, a multi-stage adaptive approach was applied in computer-based tests whereby students were assigned a block of test items based on their performance in preceding blocks.
- Test items were a mixture of multiple-choice questions and questions requiring students to construct their own responses. The items were organised into groups based on a passage of text describing a real-life situation. More than 15 hours of test items for reading, mathematics, science and global competence were covered, with different students taking different combinations of test items.
- Students also answered a background questionnaire, which took about 35 minutes to complete. The questionnaire sought information about the students themselves, their attitudes, dispositions and beliefs, their homes, and their school and learning experiences. School principals completed a questionnaire that covered school management and organisation, and the learning environment.
- Some countries/economies also distributed additional questionnaires to elicit more information. These included: in 19 countries/economies, a questionnaire for teachers asking about themselves and their teaching practices; and in 17 countries/economies, a questionnaire for parents asking them to provide information about their perceptions of and involvement in their child's school and learning.
- Countries/economies could also choose to distribute three other optional questionnaires for students: 52 countries/economies distributed a questionnaire about students' familiarity with computers; 32 countries/economies distributed a questionnaire about students' expectations for further education; and 9 countries/economies distributed a questionnaire, developed for PISA 2018, about students' well-being.

HOW IS THE ASSESSMENT CONDUCTED?

As was done in 2015, PISA 2018 delivered the assessment of all subjects via computer. Paper-based assessments were provided for countries that were not able to test their students by computer, but the paper-based assessment was limited to reading, mathematics and science trend items, which were originally developed for previous PISA assessments.[1] Since 2015, new items were developed for the computer-based assessment only.

The 2018 computer-based assessment was designed as a two-hour test. Each test form allocated to students comprised four 30-minute clusters of test material. For the main subject of reading, material equivalent to 15 30-minute clusters was developed. This material was organised into blocks instead of clusters, as the PISA 2018 reading assessment took a multi-stage adaptive approach. The reading assessment was composed of a core stage followed by stage 1 and stage 2. In stages 1 and 2, students were assigned blocks of items of either greater or lesser difficulty, depending on their performance in earlier stages (see Chapter 1 in this volume, for more detailed information on the multi-stage adaptive approach). To measure trends in the subjects of mathematics and science, six clusters were included in each subject. In addition, four clusters of global competence items were developed.[2] There were 72 different test forms.[3] Students spent one hour on the reading assessment plus one hour on one or two other subjects – mathematics, science or global competence.

Countries that used paper-based delivery for the main survey measured student performance with 30 pencil-and-paper forms containing trend items in the three core PISA subjects. The reading items in these paper-based forms were based on the 2009 reading literacy framework and did not include any items based on the new 2018 reading literacy framework.

The assessment of financial literacy was offered as an option in PISA 2018. It was based on the same framework as that developed for PISA 2012, which was also used in PISA 2015.[4] The financial literacy assessment lasted one hour (in addition to the regular PISA assessment) and comprised two clusters distributed to a subsample of students in combination with the reading and mathematics assessments.

To gather contextual information, PISA 2018 asked students and the principal of their school to respond to questionnaires. The student questionnaire took about 35 minutes to complete; the questionnaire for principals took about 45 minutes to complete. The responses to the questionnaires were analysed with the assessment results to provide both a broader and more nuanced picture of student, school and system performance. The *PISA 2018 Assessment and Analytical Framework* (OECD, 2019[1]) describes the genesis of the questionnaires in detail. The questionnaires from all assessments since PISA's inception are available on the PISA website: www.oecd.org/pisa.

The questionnaires seek information about:

- students and their family backgrounds, including their economic, social and cultural capital
- aspects of students' lives, such as their attitudes towards learning, their habits and life in and outside of school, and their family environment
- aspects of schools, such as the quality of the schools' human and material resources, public and private management and funding, decision-making processes, staffing practices, the school's curricular emphasis and the extracurricular activities it offers
- the context of instruction, including institutional structures and types, class size, classroom and school climate, and reading activities in class
- aspects of learning, including students' interest, motivation and engagement.

In PISA 2018, five additional questionnaires were offered as options:

- **computer familiarity questionnaire**, focusing on the availability and use of information and communications technologies (ICT), and on students' ability to carry out tasks on computers and their attitudes towards using computers
- **well-being questionnaire**, (new to PISA 2018) on students' perceptions of their health, life satisfaction, social connections and activities in and outside of school
- **educational career questionnaire**, which collects additional information on interruptions in schooling, preparation for students' future career, and support with language learning
- **parent questionnaire**, focusing on parents' perceptions of and involvement in their child's school, their support for learning at home, school choice, their child's career expectations, and their background (immigrant/non-immigrant)
- **teacher questionnaire**, which asks about teachers' initial training and professional development, their beliefs and attitudes, and their teaching practices. Separate questionnaires were developed for teachers of the test language and for other teachers in the school.

The contextual information collected through the student, school and optional questionnaires is complemented by system-level data. Indicators describing the general structure of each education system, such as expenditure on education, stratification, assessments and examinations, appraisals of teachers and school leaders, instruction time, teachers' salaries, actual teaching time and teacher training are routinely developed and analysed by the OECD. These data are extracted from the annual OECD publication, *Education at a Glance: OECD Indicators*, for the countries that participate in the annual OECD data collection administered through the OECD Indicators of Education Systems (INES) Network. For other countries and economies, a special system-level data collection was conducted in collaboration with PISA Governing Board members and National Project Managers.

WHO ARE THE PISA STUDENTS?

Differences between countries in the nature and extent of pre-primary education and care, the age at entry into formal schooling, the structure of the education system, and the prevalence of grade repetition mean that school grade levels are often not good indicators of where students are in their cognitive development. To better compare student performance internationally, PISA targets students of a specific age. PISA students are aged between 15 years 3 months and 16 years 2 months at the time of the assessment, and they have completed at least 6 years of formal schooling. They can be enrolled in any type of institution, participate in full-time or part-time education, in academic or vocational programmes, and attend public or private schools or foreign schools within the country. (For an operational definition of this target population, see Annex A2.) Using this age across countries and over time allows PISA to consistently compare the knowledge and skills of individuals born in the same year who are still in school at age 15, despite the diversity of their education histories in and outside of school.

The population of PISA-participating students is defined by strict technical standards, as are the students who are excluded from participating (see Annex A2). The overall exclusion rate within a country is required to be below 5% to ensure that, under reasonable assumptions, any distortions in national mean scores would remain within plus or minus 5 score points, i.e. typically within the order of magnitude of 2 standard errors of sampling. Exclusion could take place either through the schools that participated or the students who participated within schools (see Annex A2).

There are several reasons why a school or a student could be excluded from PISA. Schools might be excluded because they are situated in remote regions and are inaccessible, because they are very small, or because of organisational or operational factors that precluded participation. Students might be excluded because of intellectual disability or limited proficiency in the language of the assessment. In 31 of the 79 countries and economies that participated in PISA 2018, the percentage of school-level exclusions amounted to less than 1%; it was 4% or less in all except five countries. When the exclusion of students who met the internationally established exclusion criteria is also taken into account, the exclusion rates increase slightly. However, in 2018, the overall exclusion rate remained below 2% in 28 participating countries and economies, below 5% in 63 participating countries and economies, and below 7% in all countries except Sweden (11.1%), Israel (10.2%), Luxembourg and Norway (both 7.9%). For more detailed information about school and student exclusion from PISA 2018, see Annex A2.

WHERE CAN YOU FIND THE RESULTS?

The initial PISA 2018 results are released in six volumes:

- ***Volume I: What Students Know and Can Do*** (OECD, 2019[2]) provides a detailed examination of student performance in reading, mathematics and science, and describes how performance has changed over time.
- ***Volume II: Where All Students Can Succeed*** (OECD, 2019[3]) examines gender differences in student performance, the link between students' socio-economic status and immigrant background, on the one hand, and their performance and other outcomes, on the other, and the relationship between all of these variables and students' well-being. Trends in these indicators over time are examined when comparable data are available.
- ***Volume III: What School Life Means for Students' Lives*** (OECD, 2019[4]) focuses on the physical and emotional health of students, the role of teachers and parents in shaping the school climate, and the social life at school. The volume also examines indicators of student well-being, and how these are related to school climate.
- ***Volume IV: Are Students Smart about Money?*** (OECD, forthcoming[5]) examines 15-year-old students' understanding about money matters in the 21 countries and economies that participated in this optional assessment. The volume explores how the financial literacy of 15-year-old students is associated with their competencies in reading and mathematics, with their socio-economic status, and with their previous experiences with money. It also offers an overview of financial education in schools in the participating countries and economies, and provides case studies.
- ***Volume V: Effective Policies, Successful Schools*** (OECD, forthcoming[6]) analyses schools and school systems and their relationship with education outcomes more generally. The volume covers school governance, selecting and grouping students, and the human, financial, educational and time resources allocated to teaching and learning. Trends in these indicators are examined when comparable data are available.
- ***Volume VI: Are Students Ready to Thrive in Global Societies?*** (OECD, forthcoming[7]) examines students' ability to consider local, global and intercultural issues, understand and appreciate different perspectives and world views, interact respectfully with others, and take responsible action towards sustainability and collective well-being. It does so through both an assessment completed by students and questionnaires completed by students and school principals.[5]

Volumes II and III are published at the same time as Volume I, in December 2019; Volumes IV, V and VI are published in 2020.

The frameworks for assessing reading, mathematics, science, financial literacy and global competence in 2018 are described in the *PISA 2018 Assessment and Analytical Framework* (OECD, 2019[1]). The framework for reading is also summarised in Volume I.

Technical annexes at the end of this volume describe how questionnaire indices were constructed and discuss sampling issues, quality-assurance procedures and the process followed for developing the assessment instruments. Many of the issues covered in the technical annexes are elaborated in greater detail in the *PISA 2018 Technical Report* (OECD, forthcoming[8]).

A selection of key tables referred to in the analyses are included at the end of the respective volume in Annex B1, and a set of additional data tables is available on line (www.oecd.org/pisa). A Reader's Guide is also provided in each volume to aid in interpreting the tables and figures that accompany the report. Data from regions within the participating countries are included in Annex B2.

Notes

1. The paper-based form was used in nine countries: Argentina, Jordan, Lebanon, the Republic of Moldova, the Republic of North Macedonia, Romania, Saudi Arabia, Ukraine and Viet Nam.
2. The global competence assessment was not available in the countries/economies that conducted the PISA 2018 assessment on paper. It was conducted in Albania, Brunei Darussalam, Canada, Chile, Colombia, Costa Rica, Croatia, Greece, Hong Kong (China), Indonesia, Israel, Kazakhstan, Korea, Latvia, Lithuania, Malta, Morocco, Panama, the Philippines, the Russian Federation, Serbia, Singapore, the Slovak Republic, Spain, Chinese Taipei, Thailand and Scotland (United Kingdom). However, the global competence module was included in the student questionnaire, which was distributed in 56 of the countries/economies that took part in PISA 2018.
3. Thirty-six test forms were prepared for countries that did not participate in the global competence assessment. The number of distinct test forms is much higher when the many possible combinations of reading questions are also considered.
4. The financial literacy assessment was conducted in Australia, Brazil, Bulgaria, Canada, Chile, Estonia, Finland, Georgia, Indonesia, Italy, Latvia, Lithuania, the Netherlands, Peru, Poland, Portugal, the Russian Federation, Serbia, the Slovak Republic, Spain and the United States.
5. The global competence assessment was conducted in 27 countries and economies, while the global competence module was included in questionnaires distributed in 56 countries and economies.

References

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [1]

OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en. [2]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [3]

OECD (2019), *PISA 2018 Results (Volume III): What School Life Means for Students' Lives,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/acd78851-en. [4]

OECD (forthcoming), *PISA 2018 Results (Volume IV): Are Students Smart about Money?,* PISA, OECD Publishing, Paris. [5]

OECD (forthcoming), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools,* PISA, OECD Publishing, Paris. [6]

OECD (forthcoming), *PISA 2018 Results (Volume VI): Are Students Ready to Thrive in Global Societies?,* PISA, OECD Publishing, Paris. [7]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [8]

1

How does PISA assess reading?

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

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

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

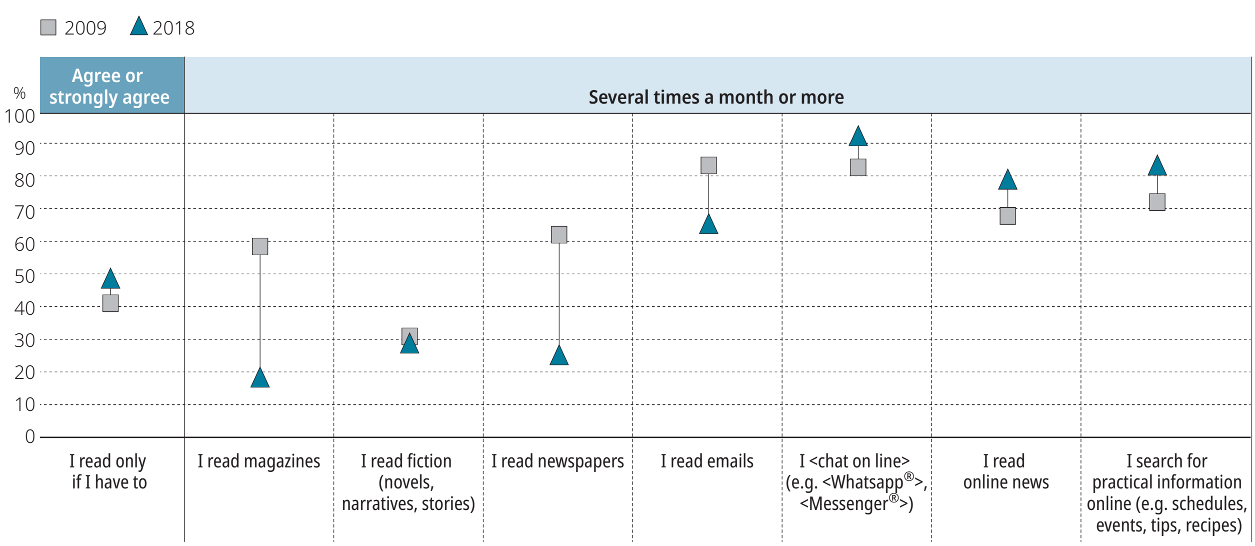
Box I.1.1. **The changing nature of reading**

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

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

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

Percentage of students; OECD average



Note: All changes between PISA 2018 and PISA 2009 are statistically significant (see Annex A3).

Source: OECD, PISA 2018 Database, Tables I.B1.57, I.B1.58 and I.B1.59.

StatLink https://doi.org/10.1787/888934028159

• • •

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

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

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

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

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

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

- a greater emphasis on multiple-source texts, i.e. texts composed of several units of text, created separately by different authors (Rouet, Britt and Potocki, 2019[2]). These types of text are more prevalent in the information-rich digital world, and the digital delivery of the PISA reading assessment made it possible to present them to students. While the availability of multiple sources does not necessarily imply greater difficulty, including multiple-source units helped to expand the range of higher-level reading processes and strategies measured by PISA. In 2018, these included searching for information across multiple documents, integrating across texts to generate inferences, assessing the quality and credibility of sources, and handling conflicts across sources (List and Alexander, 2018[3]; Barzilai, Zohar and Mor-Hagani, 2018[4]; Stadtler and Bromme, 2014[5]; Magliano et al., 2017[6])
- the explicit assessment of reading fluency, defined as the ease and efficiency with which students can read text
- the use of adaptive testing, whereby the electronic test form that a student saw depended on his or her answers to earlier questions
- the digital, on-screen delivery of text, which facilitated the first and third changes listed above. The 2009 assessment was conducted on paper while the 2018 assessment was conducted (by default) on computer.[2, 3] Students had to use navigational tools to move between passages of text, as there was often too much text to fit onto one screen.

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

HOW DOES PISA DEFINE READING LITERACY?

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

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

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

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

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

THE PISA 2018 FRAMEWORK FOR ASSESSING READING LITERACY

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

Texts

The PISA 2009 reading framework classified texts along four dimensions:

- **Medium**: Is the text delivered in print format or in electronic format?
- **Environment**: Was the text composed by an author or group of authors alone, without the participation of the reader, or was the text composed in a collaborative way with the potential contribution of the reader?
- **Text format**: Is it a piece of continuous prose, a non-continuous (usually list-like) matrix of writing, or a mixture of these two formats?[5]
- **Text type**: Why was the text written and how is it organised? Six major text types were identified:[6]
  - *descriptions* identify a tangible object and where it is located in space
  - *narrations* detail when and in what sequence events occurred
  - *expositions* explain or summarise an object or concept, and describe how objects and concepts relate to one another
  - *argumentations* try to persuade or convince the reader of the writer's viewpoint
  - *instructions* provide directions as to what to do
  - *transactions* aim to achieve a specific purpose (and are often in the form of letters or messages between two interlocutors).

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

- **Source** (related to the previous classification of "environment"): Is the text composed of a single unit (a single-source text) or of multiple units (a multiple-source text)?[7]
- **Organisational and navigational structure**: How do readers read and move through all of the text when only a certain portion can be displayed on the screen at any given time? **Static texts** have a simple, often linear organisational structure and make use of a low density of straightforward navigational tools, such as scroll bars and tabs. **Dynamic texts**, on the other hand, have a more intricate organisational structure and a higher density and complexity of navigational tools, such as a table of contents, hyperlinks to switch between segments of text, or interactive tools that allow the reader to communicate with others (as in social networks).

- **Text format** (unchanged from the previous framework): Is it a piece of continuous prose, a non-continuous (usually list-like) matrix of writing, or a mixture of these two formats?
- **Text type** (unchanged from the previous framework): Why was the text written and how is it organised?[8]

Processes

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

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

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

Note: Reading fluency is not included in the above table. Reading-fluency items were included at the beginning of the assessment and considered in the computation of students' overall score. However, these items were not included in the computation of subscale scores (neither the text-source subscale nor the reading-process subscale) and are not part of any of the percentages in this table.

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

Reading fluently

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

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

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

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

Locating information

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

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

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

- **Scanning and locating**, where readers need to scan only a single piece of text to retrieve a few words, phrases or numerical values. There is little need to comprehend the overall text as the target information appears essentially verbatim in the text.
- **Searching for and selecting relevant text**, where readers need to deal with several pieces of text. This is particularly relevant in digital reading, where the total amount of text available far exceeds the amount that readers can or need to process. In order to locate the desired information, readers need first to identify the appropriate piece of text, which adds to the complexity of this process. Text organisers, such as headers, source information (e.g. author, medium and date of publication) and links (e.g. search engine result pages) are particularly important for this process.

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

Understanding

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

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

Evaluating and reflecting

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

Three specific cognitive processes are classified under evaluating and reflecting:

- **Assessing quality and credibility**, where readers judge whether the content is valid, accurate and/or unbiased. This may also involve identifying the source of the information and thereby identifying the author's intentions and judging whether the author is competent and well-informed. Assessing quality and credibility, in other words, requires the reader to combine the content of what is said in the text with peripheral cues, such as who wrote it, when, for what purpose and so forth.
- **Reflecting on content and form**, where readers evaluate the quality and the style of the text. They need to assess whether the content and form adequately express the author's purpose and point of view. In order to do so, they may need to draw from their real-world knowledge and experience in order to be able to compare different perspectives.
- **Corroborating and handling conflict**, where readers need to compare information across texts, recognise contradictions between pieces of text and then decide how best to manage such contradictions. They can do so by evaluating the credibility of the sources, and the logic and soundness of their claims (Stadtler and Bromme, 2014[5]). This cognitive process is commonly used when examining multiple-source texts.

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

Tasks

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

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

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

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

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

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

HOW DOES THE PISA ADAPTIVE TEST OF READING WORK?

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

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

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

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

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

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

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

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

Notes

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Yamamoto, K., H. Shin and **L. Khorramdel** (2018), "Multistage Adaptive Testing Design in International Large-Scale Assessments", *Educational Measurement: Issues and Practice*, Vol. 37/4, pp. 16-27, http://dx.doi.org/10.1111/emip.12226. [19]

2

How PISA results are reported: What is a PISA score?

This chapter presents information about the methods behind the analysis of PISA data and how to interpret the score values; it does not contain results of the PISA 2018 tests. The chapter summarises the test-development and scaling procedures used to ensure that results are comparable across countries and with the results of previous PISA assessments, and explains how the score values can be interpreted.

HOW DOES PISA DEFINE A REPORTING SCALE?

This section summarises the test-development and scaling procedures used to ensure that PISA score points – the unit in which results of the PISA 2018 test are reported – are comparable across countries and with the results of previous PISA assessments. These procedures are described in greater detail in Annex A1 and in the *PISA 2018 Technical Report* (OECD, forthcoming[1]). The test-development procedures described in this section apply, in particular, to the computer-based test, which was used in the vast majority of countries/economies (70 out of 79). The differences between the paper-based test and the computer-based test are described in Annex A5.

How test questions were developed and selected

The first step in defining a reporting scale in PISA is developing a framework for each subject assessed. This framework provides a definition of what it means to be proficient in the subject;[1] delimits and organises the subject according to different dimensions (e.g. the cognitive component skills that underpin proficiency, the types of situations in which proficiency manifests itself, etc.); and identifies factors that have been found, in previous studies, to relate to proficiency in the subject. The framework also suggests the kind of test items (tasks or problems) that can be used within the constraints of the PISA design (e.g. length of the assessment, target population) to measure what students can do in the subject at different levels of proficiency (OECD, 2019[2]).

This test framework is developed by a group of international experts for each subject and is agreed upon by the participating countries. For the assessment of reading, mathematics and science, the framework is revisited every third assessment. For PISA 2018, the reading framework was redeveloped, while the mathematics and science frameworks remained identical to those used in 2015.[2] This new framework for the assessment of reading is summarised in Chapter 1 of this volume.

Once the participating countries and economies agree on the framework, the actual tasks (or items) used to assess proficiency in the subject are proposed by a consortium of testing organisations. This consortium, under contract by the OECD on behalf of participating governments, develops new items and selects items from existing tests, particularly previous PISA tests of the same subject. The expert group that developed the framework reviews the testing instruments – i.e. single items or tasks, as well as the complete electronic test forms and paper booklets – to confirm that they meet the requirements and specifications of the framework. All participating countries and economies review all of the draft items to confirm that the content, cognitive demands and contexts of the items are appropriate for a test for 15-year-olds.

It is inevitable that not all tasks in the PISA assessment are equally appropriate in different cultural contexts, and equally relevant in different curricular and instructional contexts. To address this dilemma, PISA asked experts from every participating country/economy to identify those draft tasks that they considered most appropriate for an international test. These ratings were considered when selecting items for the assessment.

Items that passed these qualitative reviews by national and international experts were translated, and these translations were carefully verified by the PISA consortium.[3] The items were then presented to a sample of 15-year-old students in all participating countries as part of a field trial to ensure that they met stringent quantitative standards of technical quality and international comparability. In particular, the field trial served to verify the psychometric equivalence of the items and test across countries, which was further examined before scaling the results of the main study (see Annex A6).

All countries that participated in the PISA 2018 assessment had to review the test material for curricular relevance, appropriateness and potential interest for 15-year-olds; and all countries were required to conduct a field trial. After the qualitative review and then again after the field trial, material was considered for rejection, revision or retention in the pool of potential items. The international expert group for each subject then formulated recommendations as to which items should be included in the main assessments. The final set of selected items was also subject to review by all countries and economies (see Annex A6). During those reviews, countries/economies provided recommendations regarding the items' suitability for assessing the competencies enumerated in the framework; the items' acceptability and appropriateness in their own national context; and the overall quality of the assessment items, all to ensure that they were of the highest standard possible. This selection was balanced across the various dimensions specified in the framework and spanned various levels of difficulty, so that the entire pool of items could measure performance across all component skills and across a broad range of contexts and student abilities. For further details, see the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

How the electronic test forms were designed

All students completed two hours of testing in two or three subjects.[4] In order to ensure that the assessment covered a wide range of content, with the understanding that each student could complete only a limited set of tasks, the full set of tasks was distributed across several different electronic test forms with overlapping content. Each student thus completed only a fraction of all items, depending on which test form he or she was assigned. This design ensures that PISA can provide valid and reliable estimates of performance at aggregate levels when considering many students together (e.g. all students in a country, or with a particular background characteristic in common).

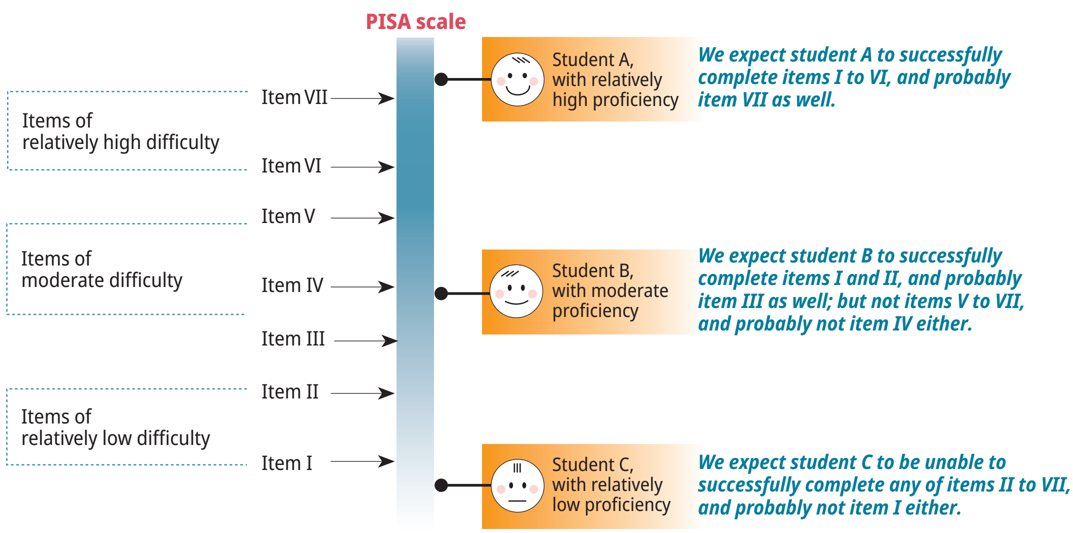
All forms contained an hour-long sequence of reading questions in the first or second part of the two-hour test, with the other hour used to assess one or sometimes two remaining subjects, which were randomly assigned. The exact sequence of test questions in reading was determined by a combination of random assignment and assignment based on performance in the initial stages of the reading assessment (see the section "How does the PISA adaptive test of reading work?" in Chapter 1). In all other subjects, the assignment of questions to students was determined by a single random draw, amongst a predetermined set of item sequences, so that each question was presented to students with equal probability and at different points during the test.

From test questions to PISA scores

PISA reports both the difficulty of questions and the proficiency of test-takers on a single continuous scale (Figure I.2.1), based on item-response theory models (see Annex A1). By showing the difficulty of each question on this scale, it is possible to locate the level of proficiency in the subject that the question demands. By showing the proficiency of test-takers on the same scale, it is possible to describe each test-taker's level of skill or literacy by the type of tasks that he or she can perform correctly most of the time.

Estimates of student proficiency are based on the kinds of tasks students are expected to perform successfully. This means that students are likely to be able to successfully answer questions located at or below the level of difficulty associated with their own position on the scale. Conversely, they are unlikely to be able to successfully answer questions above the level of difficulty associated with their position on the scale.[5]

Figure I.2.1 **Relationship between questions and student performance on a scale**



INTERPRETING DIFFERENCES IN PISA SCORES

PISA scores do not have a substantive meaning as they are not physical units, such as metres or grams. Instead, they are set in relation to the variation in results observed across all test participants. There is theoretically no minimum or maximum score in PISA; rather, the results are scaled to fit approximately normal distributions, with means around 500 score points and standard deviations around 100 score points. In statistical terms, a one-point difference on the PISA scale therefore corresponds to an effect size (Cohen's *d*) of 0.01; and a 10-point difference to an effect size of 0.10.

Determining proficiency levels for reporting and interpreting large differences in scores

To help users interpret what students' scores mean in substantive terms, PISA scales are divided into proficiency levels. For example, for PISA 2018, the range of difficulty of reading tasks is represented by eight levels of reading literacy: the simplest tasks in the assessment correspond to Level 1c; Levels 1b, 1a, 2, 3, 4, 5 and 6 correspond to increasingly more difficult tasks.

For each proficiency level identified in this way, descriptions were generated to define the kinds of knowledge and skills needed to complete those tasks successfully. Individuals who are proficient within the range of Level 1c are likely to be able to complete Level 1c tasks, but are unlikely to be able to complete tasks at higher levels. Level 6 includes tasks that pose the greatest challenge in terms of the skills needed to complete them successfully. Students with scores in this range are likely to be able to complete tasks located at this level and all the other tasks in the domain in question (see the following chapters for a detailed description of the proficiency levels in reading, mathematics and science).

Each proficiency level corresponds to a range of about 80 score points. Hence, score-point differences of 80 points can be interpreted as the difference in described skills and knowledge between successive proficiency levels.

Interpreting small differences in scores

Smaller differences in PISA scores cannot be expressed in terms of the difference in skills and knowledge between proficiency levels. However, they can still be compared with each other to conclude, for example, that the gender gap in one country is smaller than the average gender gap across OECD countries, or that the score-point difference between students with and without a tertiary-educated parent is larger than the score-point difference between students with and without an immigrant background.[6] For all differences, but particularly for small differences, it is also important to verify their "statistical significance" (see below).

In order to attach a substantive or practical meaning to differences of less than 80 points, it is tempting to compare them to some benchmark differences of recognised practical significance, expressed in the same units, such as the average achievement gain that children make from one year to the next (Bloom et al., 2008[3]). However, there is considerable uncertainty about how PISA score-point differences translate into a metric such as "years of schooling", and the empirical evidence is limited to a few countries and subjects.

There are, indeed, many difficulties involved in estimating the "typical" progress of a 15-year-old student from one year to the next or from one grade to the next in an international study such as PISA. Just as the quality of education differs across countries, so does the rate at which students progress through their schooling. A single number is unlikely to constitute a common benchmark for all countries. Furthermore, in any particular country, the observed difference between grades may be influenced by the particular grades considered. For example, the difference may depend on whether the student has transitioned from lower secondary to upper secondary school or has remained at the same level of education.

Because the PISA sample is defined by a particular age group, rather than a particular grade, in many countries, students who sit the PISA assessment are distributed across two or more grade levels. Based on this variation, past reports have estimated the average score-point difference across adjacent grades for countries in which a sizeable number of 15-year-olds are enrolled in at least two different grades. These estimates take into account some socio-economic and demographic differences that are also observed across grades. On average across countries, the difference between adjacent grades is about 40 score points. For more information see Table A1.2 in OECD (2013[4]; 2010[5]; 2007[6]).

But comparisons of performance amongst students of the same age across different grades cannot describe how much students gain, in PISA points, over a school year. Indeed, the students who are enrolled below the modal (or most common) grade for 15-year-olds differ in many ways from the students who are the same age but are enrolled in the modal grade for 15-year olds, as do those who are enrolled above the modal grade. Even analyses that account for differences in socio-economic and cultural status, gender and immigrant background can only imperfectly account for differences in motivation, aspirations, engagement and many other intangible factors that influence what students know, the grade in which they are enrolled, and how well they do on the PISA test.

Two types of studies can provide a better measure of the grade equivalence of PISA scores: longitudinal follow-up studies, where the same students who sat the PISA test are re-assessed later in their education, and cross-sectional designs that compare representative samples of students across adjacent age groups and grades.

In Germany, a longitudinal follow-up of the PISA 2003 cohort assessed the same 9th-grade students who participated in PISA one year later, when they were in the 10th grade. The comparisons showed that over this one-year period (which corresponds both to a different age and a different grade) students gained about 25 score points in the PISA mathematics test, on average, and progressed by a similar amount (21 points) in a test of science (Prenzel et al., 2006[7]).

In Canada, the Youth in Transition Study (YITS) followed the first PISA cohort, which sat the PISA 2000 test in reading, over their further study and work career. The most recent data were collected in 2009, when these young adults were 24, and included a re-assessment of their reading score. The mean score in reading amongst 24-year-olds in 2009 was 598 points, compared to a mean score of 541 points for the same young adults when they were 15 years old and in school (OECD, 2012[8]). This shows that students continue to progress in the competencies assessed in PISA beyond age 15. At the same time, it is not possible to know how this progress developed over the years (e.g. whether progress was continuous or whether more progress was made while students were still in secondary school than after they left secondary school). It must also be borne in mind that the PISA test does not measure the more specialised kinds of knowledge and skills that young adults acquire between the ages of 15 and 24.

In France, in 2012, 14-year-old students in 9th grade were assessed as part of a national extension to the PISA sample at the same time as 15-year-old students. The comparison of 14-year-old students in 9th grade (the modal grade for 14-year-old students in France) with students who were enrolled in the general academic track in 10th grade (15-year-old students) shows a 44 score-point difference in mathematics (Keskpaik and Salles, 2013[9]). This represents an upper bound on the average progression between the 9th and 10th grades in France, because some of the 14-year-olds who were included in the comparison went on to repeat 9th grade or moved to a vocational track in 10th grade, and these were likely to be amongst the lower-performing students in that group.

Because of the limited evidence about differences in PISA scores across school grades, for the same (or otherwise similar) students, and of the variability in these differences that is expected across subjects and countries, this report refrains from expressing PISA score differences in terms of an exact "years-of-schooling" equivalent. It uses the evidence from the cited studies only to establish an order of magnitude amongst differences that are statistically significant.[7]

WHEN IS A DIFFERENCE STATISTICALLY SIGNIFICANT? THREE SOURCES OF UNCERTAINTY IN COMPARISONS OF PISA SCORES

The results of the PISA assessments are estimates because they are obtained from samples of students, rather than from a census of all students, and because they are obtained using a limited set of assessment tasks, not the universe of all possible assessment tasks. A difference is called statistically significant if it is unlikely that such a difference could be observed in the estimates based on samples when, in fact, no true difference exists in the populations from which the samples are drawn.[8]

When students are sampled and assessment tasks are selected with scientific rigour, it is possible to determine the magnitude of the uncertainty associated with the estimate and to represent it as a "confidence interval", i.e. a range so defined that there is only a small probability (typically, less than 5%) for the true value to lie above its upper bound or below its lower bound. The confidence interval needs to be taken into account when making comparisons between estimates, or between an estimate and a particular benchmark value, so that differences that may arise simply due to the sampling of students and items are not interpreted as real differences in the populations. The designs of the PISA test and sample are determined with the aim of reducing, as much as possible, the statistical error associated with country-level statistics and therefore to narrow the confidence interval. Two sources of uncertainty are taken into account:

- *Sampling error:* The aim of a system-level assessment such as PISA is to generalise the results based on samples to the larger target population. The sampling methods used in PISA ensure not only that the samples are representative, and provide a valid estimate of the mean score and distribution of the population, but also that the error due to sampling is minimised, within the given budget and design constraints. The sampling error decreases the greater the number of schools and (to a lesser extent) of students included in the assessment. (In PISA, schools are the primary sampling unit, and students are sampled only from within the schools selected in the first stage of sampling.) The sampling error associated with a country's mean performance estimate is, for most countries, around two to three PISA score points. For the OECD average (which is based on 37 independent national samples) the sampling error is reduced to about 0.4 of a PISA score point.
- *Measurement error* (also called imputation error): No test is perfect or can fully measure proficiency in broad subjects such as reading, mathematics or science. The use of a limited number of items to assess proficiency in these subjects introduces some measurement uncertainty: would the use of a different set of items have resulted in different performance? This uncertainty is quantified in PISA. Amongst other things, it decreases with the number of items in a subject that underlie an estimate of proficiency. It is therefore somewhat larger for the minor subjects in an assessment than for major ones, and it is larger for individual students (who see only a fraction of all test items) than for country means (which are based on all test items). It also decreases with the amount of background information available. For estimates of country means, the imputation error is smaller than the sampling error (around 0.5 of a PISA score point in reading, and 0.8 of a point in mathematics and science).

When comparing results across different PISA assessments, an additional source of uncertainty must be taken into account. Indeed, even if different PISA assessments use the same unit for measuring performance (the metric for reading literacy, for example, was defined in PISA 2000, when reading was, for the first time, the major focus of the PISA test), the test instruments and items change in each assessment, as do the calibration samples and sometimes the statistical models used for scaling results. To make the results directly comparable over time, scales need to be equated. This means that results are transformed so that they can be expressed on the same metric. The *link error* quantifies the uncertainty around the equating of scales.

The link error represents uncertainty around the scale values ("is a score of 432 in PISA 2018 the same 432 as in PISA 2015?") and is therefore independent of the size of the student sample. As a result, it is the same for estimates based on individual countries, on subpopulations or on the OECD average.[9] For comparisons between reading results in PISA 2018 and reading results in past PISA assessments, the link error corresponds to at least 3.5 score points, making it by far the most significant source of uncertainty in trend comparisons. The link error is considerably smaller only for comparisons between PISA 2018 and PISA 2015 mathematics and science results (about 2.3 score points in mathematics and 1.5 point in science). The reduction in the uncertainty around trend comparisons is the result of improvements to the test design (in particular, a greater number of trend items common to the two assessments) and to the scaling procedure (with the introduction of concurrent calibration) introduced in PISA 2015, and of the absence of framework revisions (the frameworks for assessing mathematics and science remained unchanged from 2015). This reduced uncertainty can explain why a particular score difference may not be considered statistically significant when it is observed between PISA 2018 and PISA 2012, while a score difference of the same magnitude is considered statistically significant when it is observed between PISA 2018 and PISA 2015 (link errors for all possible score comparisons are provided in Annex A7).

Notes

1. Proficiency in reading, mathematics and science is not conceived as an attribute that a student has or does not have; rather, as an attribute that can be acquired to a greater or lesser extent.
2. The PISA 2018 paper-based instruments were based on the PISA 2009 reading framework and the PISA 2006 science framework. Only the mathematics framework was common to both the paper- and computer-based tests in 2018.
3. "Translation" also refers here to the adaptation process; see Chapter 5 in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).
4. In some countries, students with special needs received a one-hour test. This so-called "UH form" consisted of questions from the three domains of reading, mathematics and science.
5. "Unlikely", in this context, refers to a probability below 62% (see Annex A1). The farther above the student's position on the scale a question lies, the lower the probability that the student will answer successfully.
6. Comparisons of score-point differences around similar scale locations should be preferred to comparisons of gaps at different scale locations. Indeed, comparisons of gaps at different scale locations rely on equal-interval properties of the reporting scale (i.e. the idea that the difference between 300 and 350 is, in some sense, the same difference as between 700 and 750) that may not be warranted (Braun and von Davier, 2017[10]; Jacob and Rothstein, 2016[11]).
7. Woessman (2016, p. 6[12]) writes: "As a rule of thumb, learning gains on most national and international tests during one year are equal to between one-quarter and one-third of a standard deviation, which is 25-30 points on the PISA scale". This is, admittedly, a broad generalisation; without taking it too literally, this "rule of thumb" can be used to gain a sense of magnitude for score-point differences.
8. Some small countries/economies actually do conduct a census of schools and, in some cases, of students. Even in these countries/economies, PISA respondents may not coincide with the full, desired target population due to non-response and non-participation.
9. In PISA the link error is assumed to be constant across the scale. For PISA 2018 (as was the case for PISA 2015), link errors are estimated based on the variation in country means across distinct scale calibrations (see Annex A7).

References

Bloom, H. et al. (2008), "Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions", *Journal of Research on Educational Effectiveness*, Vol. 1/4, pp. 289-328, http://dx.doi.org/10.1080/19345740802400072. [3]

Braun, H. and **M. von Davier** (2017), "The use of test scores from large-scale assessment surveys: psychometric and statistical considerations", *Large-scale Assessments in Education*, Vol. 5/1, http://dx.doi.org/10.1186/s40536-017-0050-x. [10]

Jacob, B. and **J. Rothstein** (2016), "The Measurement of Student Ability in Modern Assessment Systems", *Journal of Economic Perspectives*, Vol. 30/3, pp. 85-108, http://dx.doi.org/10.1257/jep.30.3.85. [11]

Keskpaik, S. and **F. Salles** (2013), "Les élèves de 15 ans en France selon PISA 2012 en culture mathématique: baisse des performances et augmentation des inégalités depuis 2003", *Note d'information*, Vol. 13/31. [9]

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [2]

OECD (2013), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264201156-en. [4]

OECD (2012), *Learning beyond Fifteen: Ten Years after PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264172104-en. [8]

OECD (2010), *PISA 2009 Results: What Makes a School Successful?: Resources, Policies and Practices (Volume IV)*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264091559-en. [5]

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264040014-en. [6]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

Prenzel, M. et al. (eds.) (2006), *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*, Waxmann Verlag GmbH. [7]

Woessmann, L. (2016), "The Importance of School Systems: Evidence from International Differences in Student Achievement", *Journal of Economic Perspectives*, Vol. 30/3, pp. 3-32, http://dx.doi.org/10.1257/jep.30.3.3. [12]

3

Who sits the PISA assessment?

This chapter describes the students in the PISA target population, or those who were eligible to sit the PISA assessment. It discusses the extent to which this target population represented (or covered) the total population. The chapter also presents the grade distribution of the students who participated in the assessment.

What the data tell us

- Over 97% of 15-year-olds in Brunei Darussalam, Germany, Hong Kong (China), Malta and Slovenia were represented in the PISA sample. However, less than 80% of the total population of 15-year-olds were represented in 19 of the 79 PISA-participating education systems, of which 3 were OECD countries.
- In most PISA-participating education systems, 15-year-old students were most commonly enrolled in grade 10 at the time of the assessment.

WHO IS THE PISA TARGET POPULATION?

PISA 2018 assessed the cumulative outcomes of education and learning at a point at which most children are still enrolled in formal education: the age of 15. In particular, PISA assessed students who were aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months[1] at the time of the assessment.[2] This was done so that students could be compared across countries shortly before they are faced with decisions about major life choices, such as entering the workforce or pursuing further education. Students at the same grade level could have been selected instead, but differences in the institutional nature of education systems (e.g. the age at entry into pre-primary school and into formal schooling; grade-retention policies; and even whether the meaning of grades is equivalent across countries) make it more difficult to make a fair comparison about how prepared students are for life post-school.[3]

The 15-year-olds in the PISA sample must also have been enrolled in an educational institution[4] at grade 7 or higher. All such students were eligible to sit the PISA assessment, regardless of the type of educational establishment in which they were enrolled and whether they were enrolled in full-time or part-time education.

HOW MANY 15-YEAR-OLDS DOES THE PISA SAMPLE REPRESENT?

Not all of the students who were eligible to sit the PISA assessment were actually assessed. A two-stage sampling procedure first selected a representative sample of at least 150 schools, taking into account factors such as location (state or province; but also whether the school is located in a rural area, town or city) and level of education (lower secondary or upper secondary school). Then, in the second stage, roughly 42 15-year-old students were selected from each school to sit the assessment.[5] In PISA 2018, most countries assessed between 4 000 and 8 000 students.[6]

Students selected to sit the PISA assessment were assigned sampling weights so as to represent the entire PISA-eligible cohort. However, some otherwise-eligible 15-year-old students enrolled in grade 7 or above could be excluded for various reasons, including the remoteness and inaccessibility of their school, intellectual or physical disability, a lack of proficiency in the test language, or a lack of test material in the language of instruction.

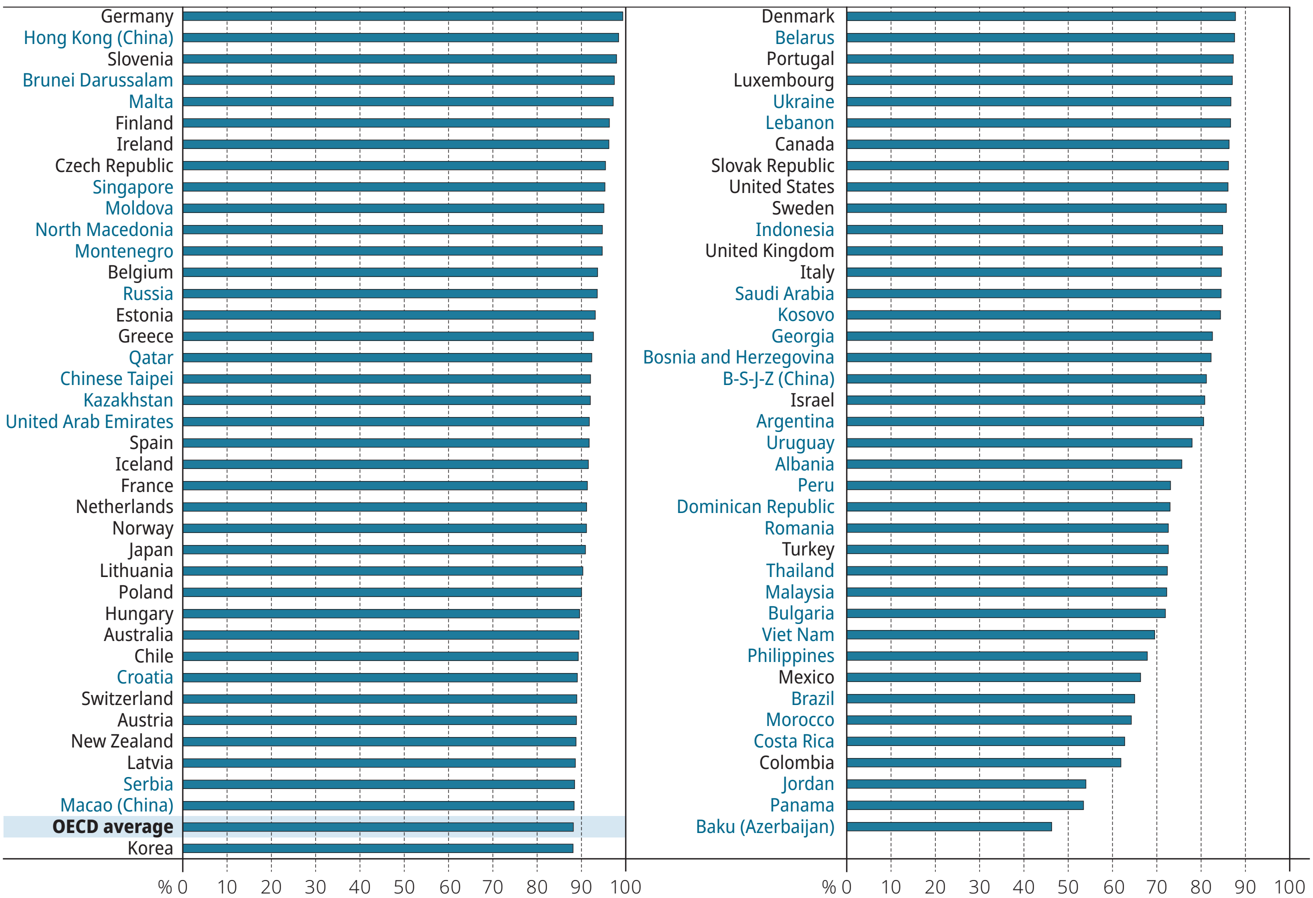
Figure I.3.1 shows the proportion of 15-year-olds in each country/economy who were covered by the PISA sample, also known as Coverage Index 3. It ranged from over 99% in Germany, over 98% in Hong Kong (China), and over 97% in Brunei Darussalam, Malta and Slovenia, to under 50% in Baku (Azerbaijan) and under 60% in Jordan and Panama. This proportion exceeds 80% in most OECD countries; only Colombia (62%), Mexico (66%) and Turkey (73%) did not reach this level (Table I.A2.1).[7]

In most countries, low values of Coverage Index 3, which is based on verified (rather than on reported) enrolment, can be attributed to 15-year-olds who were no longer enrolled in school or who had been held back in primary school. Coverage Index 3 may also have been lower due to student exclusions from the test and dropout during the school year. For example, in Colombia, official enrolment rates indicate that 75% of 15-year-olds were enrolled in grade 7 or above while Coverage Index 3 indicates that only 62% of 15-year-olds were eligible to sit the PISA test.

The exclusion of students and schools generally has a limited impact on Coverage Index 3. The overall exclusion rate of students enrolled in school was less than 1% in 14 of the 79 education systems that participated in PISA 2018 (of which 11 were partner countries and economies); this rate exceeded 5% in only 16 education systems (of which 14 were OECD countries) (Figure I.3.2 and Table I.A2.1). Further information was sought for all countries/economies in which the exclusion rate exceeded 5%. In all cases, it was found that there were no significant biases to the results in reading, mathematics and science because over 5% of students had been excluded and the data were deemed to be acceptable. Results for these education systems are still comparable across education systems and across PISA cycles. Please see Annex A2 for country-specific details.

Figure I.3.1 **Percentage of 15-year-olds covered by PISA**

Coverage Index 3



Countries and economies are ranked in descending order of coverage.

Source: OECD, PISA 2018 Database, Table I.A2.1.

StatLink https://doi.org/10.1787/888934028178

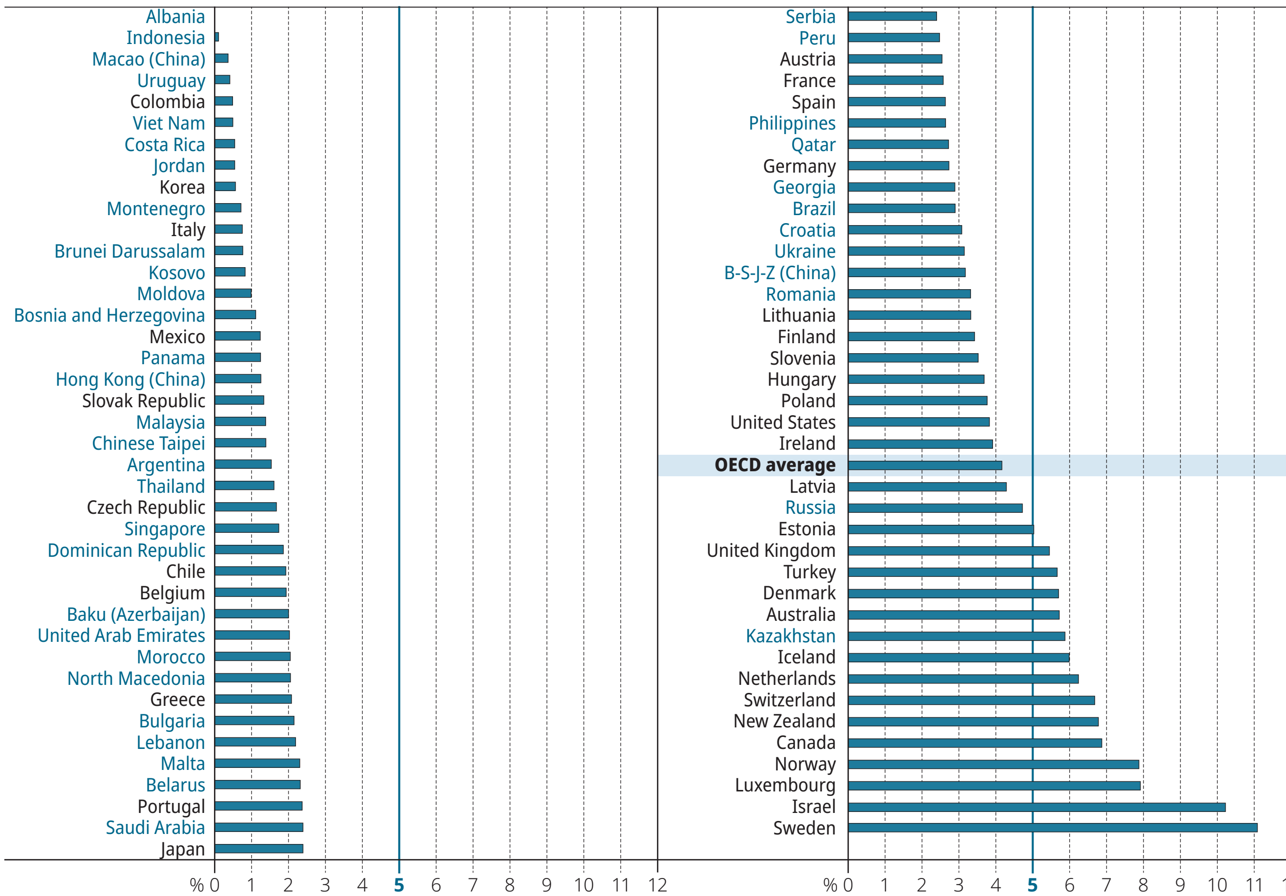
Access to schooling is a prerequisite for achieving equity in education. As discussed in Chapter 1 of *PISA 2018 Results (Volume II): Where All Students Can Succeed* (2019[1]), PISA 2018 considered two dimensions of equity: inclusion and fairness. PISA defines inclusion in education as ensuring that all students attain essential foundational skills; it relates fairness to the distribution of opportunities to acquire a quality education and, more specifically, to the degree to which background circumstances are related to students' education outcomes. While enrolling all 15-year-olds in school does not guarantee that every student will acquire the skills needed to thrive in an increasingly knowledge-intensive economy, it is the first step towards building a more inclusive and fairer education system.

Students who have already left formal schooling by the age of 15 are more likely to be cognitively weaker than those who remain at school (Spaull and Taylor, 2015[2]; Taylor and Spaull, 2015[3]; Hanushek and Woessmann, 2008[4]). Therefore, the performance of the 15-year-old cohort in an education system is likely to be overestimated if PISA results are not considered in the context of coverage, and the extent of this overestimation is likely to increase as the coverage decreases. Chapter 9 relates changes in PISA scores in a country/economy with changes in the coverage of its 15-year-old cohort.

Coverage of the 15-year-old cohort has increased in many countries since earlier PISA assessments (see Chapter 9 and Table I.A2.2). Between 2003 and 2018, Indonesia added almost 1.8 million students, and Mexico and Turkey more than 400 000 students, to the total population of 15-year-olds eligible to participate in PISA. Uruguay and Brazil also increased the number of 15-year-olds eligible to participate in PISA, despite having a smaller population of 15-year-olds in 2018 than in 2003. As a result, PISA coverage – the number obtained by dividing the number of PISA-eligible students by the total number of 15-year-olds in a country – increased greatly in all five countries as well as in Albania, Costa Rica and Lebanon.[8] This might reflect the expansion of schooling to previously underserved communities, especially in developing countries, and could be indicative of progress towards the United Nations Sustainable Development Goal for quality education in these countries (see Chapter 10).[9]

Figure I.3.2 **Overall exclusion rate from the PISA sample**

Percentage of the national desired target population (15-year-olds enrolled in school in grade 7 or above) excluded from the PISA sample (through either school-level or in-school exclusions)



Note: The vertical blue line corresponds to the 5% exclusion limit. All countries/economies that exceeded this threshold were required to provide documentation to ensure that the exclusions did not affect the comparability of their results with those from other countries/economies.

Countries and economies are ranked in ascending order of the overall exclusion rate.

Source: OECD, PISA 2018 Database, Table I.A2.1.

StatLink https://doi.org/10.1787/888934028197

However, a large decrease in coverage was observed in some countries too, particularly in Jordan (by about 20 percentage points since 2006 and 2009, when it first participated in PISA). In Jordan, the population of 15-year-olds represented by PISA increased by about 25 000 since 2006, but the total population of 15-year-olds grew by about 90 000, largely as a result of the massive influx of refugees from neighbouring countries.[10] The influx of immigrant and refugee students into Sweden since 2015 also resulted in a marked increase in exclusions (by 5 percentage points) and a large drop in Coverage Index 3 (by 8 percentage points).

THE DISTRIBUTION OF PISA STUDENTS ACROSS GRADES

As alluded to above, students in different countries start formal schooling at different ages. In addition, differences in grade-retention policies and inconsistencies in school attendance mean that students can progress through school differently. Students in some countries automatically move on to the next grade each year regardless of performance, while students in other countries can be held back to repeat a year or may simply not attend school for a year or more, delaying their progress through the school system.[11]

Hence, 15-year-old students show different distributions across grade levels in different countries. In Brazil, Malta, New Zealand and the United Kingdom, the modal grade, or the grade in which 15-year-old students are most commonly found, is grade 11 (Table I.3.1); in the latter three countries, roughly 90% (or more) of students are enrolled in grade 11. These are countries in which students enter primary school at an earlier age. Grade 9 is the modal grade in 21 countries and economies – often the countries where students start formal schooling at a later age. Grade 10 is the modal grade in the remaining 53 PISA-participating countries and economies.

Table I.3.1 **Modal grade of students in the PISA sample**

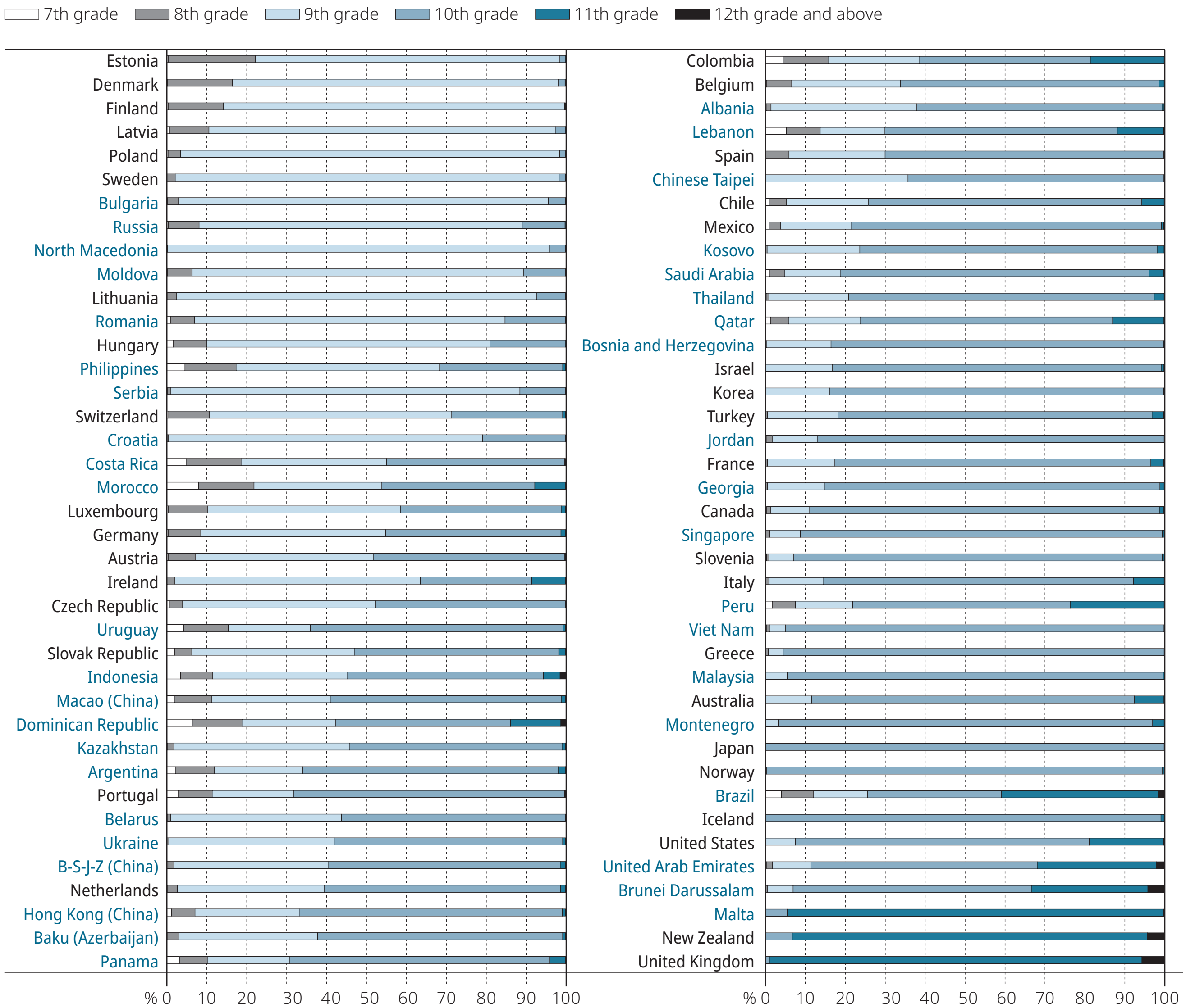
The modal grade is ...			
Grade 9	Grade 10		Grade 11
Bulgaria	Albania	Korea	Brazil
Croatia	Argentina	Kosovo	Malta
Czech Republic	Australia	Lebanon	New Zealand
Denmark	Austria	Macao (China)	United Kingdom
Estonia	Baku (Azerbaijan)	Malaysia	
Finland	Belarus	Mexico	
Germany	Belgium	Montenegro	
Hungary	Bosnia and Herzegovina	Morocco	
Ireland	Brunei Darussalam	Netherlands	
Latvia	B-S-J-Z (China)	Norway	
Lithuania	Canada	Panama	
Luxembourg	Chile	Peru	
Moldova	Colombia	Portugal	
North Macedonia	Costa Rica	Qatar	
Philippines	Cyprus	Saudi Arabia	
Poland	Dominican Republic	Singapore	
Romania	France	Slovak Republic	
Russia	Georgia	Slovenia	
Serbia	Greece	Spain	
Sweden	Hong Kong (China)	Chinese Taipei	
Switzerland	Iceland	Thailand	
	Indonesia	Turkey	
	Israel	Ukraine	
	Italy	United Arab Emirates	
	Japan	United States	
	Jordan	Uruguay	
	Kazakhstan	Viet Nam	

Source: OECD, PISA 2018 Database.

Almost 100% of students in three countries, Iceland, Japan and Norway, were enrolled in grade 10 at the time they sat the PISA test, which reflects the lack of grade retention and grade advancement in these countries, and the alignment between the PISA testing period and the cut-off dates for school enrolment (Figure I.3.3). In contrast, students in Brazil, Colombia, the Dominican Republic and Morocco were dispersed over a wide range of grade levels. Dispersion over two consecutive grade levels may be due to a misalignment between the PISA testing period and the cut-off date for school enrolment, or to a certain degree of flexibility over when children can enter formal schooling. However, in many of these countries, the wide range of distribution observed in the aforementioned education systems often reflects inconsistencies in school attendance. In these (mostly developing) countries, some students may miss a year of school because they are unable to pay school fees; when they return to school, they will necessarily be one (or more) grades behind students whose education was not interrupted. Similarly, their parents may keep them home occasionally to help out with other tasks. In these cases, students' performance might be judged to be insufficient to move on to the next grade if the students had missed too many days of school, and they may be required to repeat their current grade. By the age of 15, such students may be found in a wide range of grade levels.

By using age, rather than grade level, as the criterion for students' eligibility to sit the assessment, PISA enables a standardised comparison of the skills of students who are soon to enter adult life. However, these students may be at different points in their educational career – both between and within countries. PISA cannot capture students' progress through further education after the age of 15; students who appear to be behind today, based on their performance in the PISA assessment, may well catch up with their peers in the future.

Figure I.3.3 **Grade distribution of students in the PISA sample**



Countries and economies are ranked in ascending order of the average grade level of their 15-year-old student population.

Source: OECD, PISA 2018 Database, Table I.A2.8.

StatLink https://doi.org/10.1787/888934028216

Notes

1. More precisely, PISA assessed students who were at least 15 years and 3 (complete) months old and who were at most 16 years and 3 (complete) months old (i.e. younger than 16 years, 2 months and roughly 30 days old), with a tolerance of one month on each side of this age window. If, as was the case in most countries, the PISA assessment was conducted in April 2018, all students born in 2002 would have been eligible. For simplicity, students aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months are, in this report, referred to as being 15 years old.
2. The month of birth of a cohort that is eligible to sit the PISA test varies over the range of an entire year.
3. Fifteen-year-old students in different countries will have had different numbers of years of formal schooling. However, it is assumed that, in countries where entry into formal schooling takes place at an older age, children are still educated before they enter formal schooling, either at home or in various pre-school programmes. This should mitigate, to some extent, the disadvantage that students in these countries may face by not having attended formal schooling for as many years as students in other countries. This also contrasts with other large-scale international assessments that select students by grade level in order to measure the extent to which students master certain aspects of the curriculum. In those assessments, students in countries with a later age at entry into formal schooling are older and will typically have an advantage over students in other countries.
4. Educational institutions are generally referred to as schools in this publication, although some educational institutions (in particular, some types of vocational education establishments) may not be referred to as schools in certain countries.
5. In schools with fewer than 42 15-year-old students, all 15-year-old students in the school were selected. Countries where schools tended to have fewer than 42 students per age group could opt for a sampling design with a smaller number of students sampled per school (e.g. 30), but had to increase the overall number of schools sampled proportionately.
6. Larger numbers of students sat the PISA assessment in countries whose subnational regions were also extensively sampled. This was the case, for example, in Canada (over 22 000 participating students) and Spain (almost 36 000 participating students). In smaller education systems, almost all eligible students and schools were sampled in order to meet the criterion for sample size.
7. Data for an education system come from multiple sources that might not be consistent with one another, introducing error into the calculation of coverage indices. The total population of 15-year-olds is typically based on demographic projections from census and registry data; the total population of 15-year-old students may be based on central school registries or on household surveys; and the national desired target population (the PISA-eligible students that the PISA sample aims to represent) is based on school-enrolment estimates (for all schools) provided by the national PISA centre and on student lists provided by the sampled schools. As one example of this inconsistency, the data paradoxically show that in Brunei Darussalam, Canada, Malta and Slovenia, more students were eligible to sit the PISA assessment than were 15-year-old students in grade 7 and above, according to central school registries or household surveys.
8. Coverage Index 3 also increased by about 26 percentage points in Argentina between 2015 and 2018. The low values of Coverage Index 3 in 2015 are due to a statistical anomaly (incomplete school sampling frames); as a result, Argentina's results for 2015 are not comparable to those of other countries or to results for Argentina from previous years or for 2018.
9. Improvements in data-collection procedures may also have led to changes in the reported proportion of 15-year-olds covered by the PISA sample, although the direction of the impact of such improvements is unclear.
10. The decrease in the coverage of 15-year-olds in Jordan is likely a result of the ongoing refugee crises in neighbouring countries. Refugee children may be enrolled outside of the country's formal education system, in which case they would not be sampled by PISA.
11. In addition to grade-retention policies, certain education systems permit accelerated advancement through the school system.

References

Hanushek, E. and **L. Woessmann** (2008), "The Role of Cognitive Skills in Economic Development", *Journal of Economic Literature*, Vol. 46/3, pp. 607-668, http://dx.doi.org/10.1257/jel.46.3.607. [4]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [1]

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. [2]

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. [3]

PISA

4

How did countries perform in PISA 2018?

This chapter compares students' mean scores and the variation in their performance in reading, mathematics and science across the countries and economies that participated in the PISA 2018 assessment. It also highlights differences in social and economic contexts across education systems.

PISA outcomes are reported in a variety of ways; but the easiest way to gain an understanding of the overall performance of a country or economy is through the mean performance of its students. Because countries' and economies' standing in comparison with other countries/economies that participated in PISA can differ across subjects, this chapter includes multiple comparisons of mean performance. Further comparisons can consider the proportion of students who achieve a certain level of performance (see Chapters 5, 6 and 7 in this volume), or the extent to which learning outcomes vary within countries (see the section on "variation in performance" below and Volume II of the *PISA 2018 Results* report, *Where All Students Can Succeed* [OECD, 2019[11]]). No single ranking does justice to the richness of information that PISA provides and, more important, to the variety of goals that education systems pursue. This chapter also highlights the statistical uncertainty in PISA results when comparing countries and economies.

When considering differences in performance across countries and economies, it is also important to consider differences in context – such as a country's level of development or the proportion of 15-year-olds who are in school and eligible to sit the PISA test. These factors are discussed at the end of the chapter.

What the data tell us

- On average, students in Beijing, Shanghai, Jiangsu and Zhejiang (China) and Singapore outperformed students from all other countries in reading, mathematics and science.
- Differences in performance between students within the same country are, in general, larger than between-country differences in performance. For example, in every country and economy, the performance gap between the highest-scoring 5% of students and the lowest-scoring 5% of students in reading is larger than the difference in mean performance between the highest-performing country and the lowest-performing country.
- While an inadequately resourced education system cannot deliver good results, Estonia, with a level of expenditure on education that is about 30% lower than the OECD average, is nevertheless one of the top-performing OECD countries in reading, mathematics and science.

MEAN PERFORMANCE IN READING, MATHEMATICS AND SCIENCE

In 2018, the mean reading score amongst OECD countries was 487 points; the mean score in mathematics and science was 489 points. In reading, Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]") (555 points) and Singapore (549 points) scored significantly higher than all other countries/economies that participated in PISA 2018. In mathematics and science, the highest mean performance was achieved by students in B-S-J-Z (China) (591 points in mathematics and 590 points in science), and the second-highest mean performance by students in Singapore (569 points in mathematics and 551 points in science).

Table I.4.1, Table I.4.2, and Table I.4.3 show each country's/economy's mean score, and indicate for which pairs of countries/economies the differences between the means are statistically significant. Indeed, when comparing mean performance across countries/economies, only those differences that are statistically significant should be considered (see Chapter 2). For each country/economy shown in the middle column, the countries/economies whose mean scores are not statistically significantly different are listed in the right column. For example, B-S-J-Z (China) scored higher than Singapore on the PISA mathematics and science scales, but in reading, the mean performance of B-S-J-Z (China) was not statistically significantly different from that of Singapore; or students in Germany performed better in science than students in France, but in reading and mathematics, their mean scores were not statistically significantly different.

In Table I.4.1, Table I.4.2, and Table I.4.3, countries and economies are divided into three broad groups: those whose mean scores are statistically around the OECD mean (highlighted in white); those whose mean scores are above the OECD mean (highlighted in blue); and those whose mean scores are below the OECD mean (highlighted in grey).[1]

Twenty countries and economies performed above the OECD average in all three domains (reading, mathematics and science). B-S-J-Z (China) and Singapore were the highest-performing education systems: in all three subjects, their mean scores lay more than 50 points above the average score across OECD countries. In reading, Estonia, Canada, Finland and Ireland were the highest-performing OECD countries (the mean performance of Korea was significantly below that of Estonia, but not below those of Canada, Finland and Ireland; and Poland's score was below those of Estonia, Canada and Finland, but not below that of Ireland) (all countries/economies are listed in descending order of their mean scores).

Table I.4.1 [1/2] Comparing countries' and economies' performance in 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 and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
555	B-S-J-Z (China)	Singapore
549	Singapore	B-S-J-Z (China)
525	Macao (China)	Hong Kong (China),[1] Estonia, Finland
524	Hong Kong (China)[1]	Macao (China), Estonia, Canada, Finland, Ireland
523	Estonia	Macao (China), Hong Kong (China),[1] Canada, Finland, Ireland
520	Canada	Hong Kong (China),[1] Estonia, Finland, Ireland, Korea
520	Finland	Macao (China), Hong Kong (China),[1] Estonia, Canada, Ireland, Korea
518	Ireland	Hong Kong (China),[1] Estonia, Canada, Finland, Korea, Poland
514	Korea	Canada, Finland, Ireland, Poland, Sweden, United States[1]
512	Poland	Ireland, Korea, Sweden, New Zealand, United States[1]
506	Sweden	Korea, Poland, New Zealand, United States,[1] United Kingdom, Japan, Australia, Chinese Taipei, Denmark, Norway, Germany
506	New Zealand	Poland, Sweden, United States,[1] United Kingdom, Japan, Australia, Chinese Taipei, Denmark
505	United States[1]	Korea, Poland, Sweden, New Zealand, United Kingdom, Japan, Australia, Chinese Taipei, Denmark, Norway, Germany
504	United Kingdom	Sweden, New Zealand, United States,[1] Japan, Australia, Chinese Taipei, Denmark, Norway, Germany
504	Japan	Sweden, New Zealand, United States,[1] United Kingdom, Australia, Chinese Taipei, Denmark, Norway, Germany
503	Australia	Sweden, New Zealand, United States,[1] United Kingdom, Japan, Chinese Taipei, Denmark, Norway, Germany
503	Chinese Taipei	Sweden, New Zealand, United States,[1] United Kingdom, Japan, Australia, Denmark, Norway, Germany
501	Denmark	Sweden, New Zealand, United States,[1] United Kingdom, Japan, Australia, Chinese Taipei, Norway, Germany
499	Norway	Sweden, United States,[1] United Kingdom, Japan, Australia, Chinese Taipei, Denmark, Germany, Slovenia
498	Germany	Sweden, United States,[1] United Kingdom, Japan, Australia, Chinese Taipei, Denmark, Norway, Slovenia, Belgium, France, Portugal[1]
495	Slovenia	Norway, Germany, Belgium, France, Portugal,[1] Czech Republic
493	Belgium	Germany, Slovenia, France, Portugal,[1] Czech Republic
493	France	Germany, Slovenia, Belgium, Portugal,[1] Czech Republic
492	Portugal[1]	Germany, Slovenia, Belgium, France, Czech Republic, Netherlands[1]
490	Czech Republic	Slovenia, Belgium, France, Portugal,[1] Netherlands,[1] Austria, Switzerland
485	Netherlands[1]	Portugal,[1] Czech Republic, Austria, Switzerland, Croatia, Latvia, Russia
484	Austria	Czech Republic, Netherlands,[1] Switzerland, Croatia, Latvia, Russia
484	Switzerland	Czech Republic, Netherlands,[1] Austria, Croatia, Latvia, Russia, Italy
479	Croatia	Netherlands,[1] Austria, Switzerland, Latvia, Russia, Italy, Hungary, Lithuania, Iceland, Belarus, Israel
479	Latvia	Netherlands,[1] Austria, Switzerland, Croatia, Russia, Italy, Hungary, Lithuania, Belarus
479	Russia	Netherlands,[1] Austria, Switzerland, Croatia, Latvia, Italy, Hungary, Lithuania, Iceland, Belarus, Israel
476	Italy	Switzerland, Croatia, Latvia, Russia, Hungary, Lithuania, Iceland, Belarus, Israel
476	Hungary	Croatia, Latvia, Russia, Italy, Lithuania, Iceland, Belarus, Israel
476	Lithuania	Croatia, Latvia, Russia, Italy, Hungary, Iceland, Belarus, Israel
474	Iceland	Croatia, Russia, Italy, Hungary, Lithuania, Belarus, Israel, Luxembourg
474	Belarus	Croatia, Latvia, Russia, Italy, Hungary, Lithuania, Iceland, Israel, Luxembourg, Ukraine
470	Israel	Croatia, Russia, Italy, Hungary, Lithuania, Iceland, Belarus, Luxembourg, Ukraine, Turkey
470	Luxembourg	Iceland, Belarus, Israel, Ukraine, Turkey
466	Ukraine	Belarus, Israel, Luxembourg, Turkey, Slovak Republic, Greece
466	Turkey	Israel, Luxembourg, Ukraine, Greece
458	Slovak Republic	Ukraine, Greece, Chile
457	Greece	Ukraine, Turkey, Slovak Republic, Chile
452	Chile	Slovak Republic, Greece, Malta
448	Malta	Chile
439	Serbia	United Arab Emirates, Romania
432	United Arab Emirates	Serbia, Romania, Uruguay, Costa Rica
428	Romania	Serbia, United Arab Emirates, Uruguay, Costa Rica, Cyprus, Moldova, Montenegro, Mexico, Bulgaria, Jordan
427	Uruguay	United Arab Emirates, Romania, Costa Rica, Cyprus, Moldova, Mexico, Bulgaria
426	Costa Rica	United Arab Emirates, Romania, Uruguay, Cyprus, Moldova, Montenegro, Mexico, Bulgaria, Jordan
424	Cyprus	Romania, Uruguay, Costa Rica, Moldova, Montenegro, Mexico, Bulgaria, Jordan
424	Moldova	Romania, Uruguay, Costa Rica, Cyprus, Montenegro, Mexico, Bulgaria, Jordan
421	Montenegro	Romania, Costa Rica, Cyprus, Moldova, Mexico, Bulgaria, Jordan
420	Mexico	Romania, Uruguay, Costa Rica, Cyprus, Moldova, Montenegro, Bulgaria, Jordan, Malaysia, Colombia

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.4.

StatLink https://doi.org/10.1787/888934028235

...

Table I.4.1 [2/2] **Comparing countries' and economies' performance in reading**

(blue)	Statistically significantly **above** the OECD average
(white)	**Not statistically significantly different** from the OECD average
(grey)	Statistically significantly **below** the OECD average

Mean score	Comparison country/economy	Countries and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
420	Bulgaria	Romania, Uruguay, Costa Rica, Cyprus, Moldova, Montenegro, Mexico, Jordan, Malaysia, Brazil, Colombia
419	Jordan	Romania, Costa Rica, Cyprus, Moldova, Montenegro, Mexico, Bulgaria, Malaysia, Brazil, Colombia
415	Malaysia	Mexico, Bulgaria, Jordan, Brazil, Colombia
413	Brazil	Bulgaria, Jordan, Malaysia, Colombia
412	Colombia	Mexico, Bulgaria, Jordan, Malaysia, Brazil, Brunei Darussalam, Qatar, Albania
408	Brunei Darussalam	Colombia, Qatar, Albania, Bosnia and Herzegovina
407	Qatar	Colombia, Brunei Darussalam, Albania, Bosnia and Herzegovina, Argentina
405	Albania	Colombia, Brunei Darussalam, Qatar, Bosnia and Herzegovina, Argentina, Peru, Saudi Arabia
403	Bosnia and Herzegovina	Brunei Darussalam, Qatar, Albania, Argentina, Peru, Saudi Arabia
402	Argentina	Qatar, Albania, Bosnia and Herzegovina, Peru, Saudi Arabia
401	Peru	Albania, Bosnia and Herzegovina, Argentina, Saudi Arabia, Thailand
399	Saudi Arabia	Albania, Bosnia and Herzegovina, Argentina, Peru, Thailand
393	Thailand	Peru, Saudi Arabia, North Macedonia, Baku (Azerbaijan), Kazakhstan
393	North Macedonia	Thailand, Baku (Azerbaijan)
389	Baku (Azerbaijan)	Thailand, North Macedonia, Kazakhstan
387	Kazakhstan	Thailand, Baku (Azerbaijan)
380	Georgia	Panama
377	Panama	Georgia, Indonesia
371	Indonesia	Panama
359	Morocco	Lebanon, Kosovo
353	Lebanon	Morocco, Kosovo
353	Kosovo	Morocco, Lebanon
342	Dominican Republic	Philippines
340	Philippines	Dominican Republic

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.4.

StatLink https://doi.org/10.1787/888934028235

In science, the highest-performing OECD countries were Japan and Estonia. In mathematics, the highest-performing OECD countries were Japan, Korea and Estonia. B-S-J-Z (China), Singapore, Estonia, Canada, Finland, Ireland, Japan and Korea scored above the OECD average in all three subjects, as did Macao (China), Hong Kong (China), Chinese Taipei, Sweden, New Zealand, the United Kingdom, Denmark, Germany, Slovenia, Belgium and France (in descending order of mean performance in reading).

Two countries (the United States and Australia) scored above the OECD average in reading and science, but not in mathematics; in the United States, performance in mathematics was significantly below the OECD average, while the performance of students in Australia was not statistically significantly different from the OECD average. Norway scored above the OECD average in reading and mathematics, but close to the OECD average in science. Three countries (the Czech Republic, the Netherlands and Switzerland) scored above the OECD average in mathematics and science, but close to the OECD average in reading. Some countries achieved above-average results in one subject only; this was the case of Austria, Iceland and Latvia in mathematics.

Eight countries whose mean scores lay below the OECD average (Argentina, Jordan, Lebanon, the Republic of Moldova, the Republic of North Macedonia, Romania, Saudi Arabia and Ukraine) conducted the PISA 2018 test using pen-and-paper forms, designed initially for the PISA 2012 or earlier assessments. Their results are reported on the same scale as those of the remaining countries, just as PISA 2018 results for all remaining countries/economies are reported on the same scale as past PISA results.[2]

The gap in performance between the highest- and lowest-performing OECD countries was 111 score points in reading; it was even larger in mathematics and science.[3] But the difference between the highest-performing and lowest-performing education systems that took part in PISA 2018 was about twice as large (Table I.4.1, Table I.4.2, and Table I.4.3), and the gap in mean performance, across all education systems in the world, is likely to be even larger. Indeed, the developing countries that participated in PISA – either as part of PISA 2018 or, in 2017, as part of the PISA for Development initiative (see Chapter 11 and Ward [2018[2]]) – represent only a minority of all developing countries. They often participated with the clear understanding that their students were not learning at adequate levels, even when they were in school. By participating in a global assessment of learning outcomes, these developing countries demonstrated a strong commitment to develop an evidence base for future education reforms and to address the international "learning crisis" (World Bank, 2017[3]).

Table I.4.2 [1/2] Comparing countries' and economies' performance in 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 and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
591	B-S-J-Z (China)	
569	Singapore	
558	Macao (China)	Hong Kong (China)[1]
551	Hong Kong (China)[1]	Macao (China)
531	Chinese Taipei	Japan, Korea
527	Japan	Chinese Taipei, Korea, Estonia
526	Korea	Chinese Taipei, Japan, Estonia, Netherlands[1]
523	Estonia	Japan, Korea, Netherlands[1]
519	Netherlands[1]	Korea, Estonia, Poland, Switzerland
516	Poland	Netherlands,[1] Switzerland, Canada
515	Switzerland	Netherlands,[1] Poland, Canada, Denmark
512	Canada	Poland, Switzerland, Denmark, Slovenia, Belgium, Finland
509	Denmark	Switzerland, Canada, Slovenia, Belgium, Finland
509	Slovenia	Canada, Denmark, Belgium, Finland
508	Belgium	Canada, Denmark, Slovenia, Finland, Sweden, United Kingdom
507	Finland	Canada, Denmark, Slovenia, Belgium, Sweden, United Kingdom
502	Sweden	Belgium, Finland, United Kingdom, Norway, Germany, Ireland, Czech Republic, Austria, Latvia
502	United Kingdom	Belgium, Finland, Sweden, Norway, Germany, Ireland, Czech Republic, Austria, Latvia, France
501	Norway	Sweden, United Kingdom, Germany, Ireland, Czech Republic, Austria, Latvia, France, Iceland
500	Germany	Sweden, United Kingdom, Norway, Ireland, Czech Republic, Austria, Latvia, France, Iceland, New Zealand
500	Ireland	Sweden, United Kingdom, Norway, Germany, Czech Republic, Austria, Latvia, France, Iceland, New Zealand
499	Czech Republic	Sweden, United Kingdom, Norway, Germany, Ireland, Austria, Latvia, France, Iceland, New Zealand, Portugal[1]
499	Austria	Sweden, United Kingdom, Norway, Germany, Ireland, Czech Republic, Latvia, France, Iceland, New Zealand, Portugal[1]
496	Latvia	Sweden, United Kingdom, Norway, Germany, Ireland, Czech Republic, Austria, France, Iceland, New Zealand, Portugal,[1] Australia
495	France	United Kingdom, Norway, Germany, Ireland, Czech Republic, Austria, Latvia, Iceland, New Zealand, Portugal,[1] Australia
495	Iceland	Norway, Germany, Ireland, Czech Republic, Austria, Latvia, France, New Zealand, Portugal,[1] Australia
494	New Zealand	Germany, Ireland, Czech Republic, Austria, Latvia, France, Iceland, Portugal,[1] Australia
492	Portugal[1]	Czech Republic, Austria, Latvia, France, Iceland, New Zealand, Australia, Russia, Italy, Slovak Republic
491	Australia	Latvia, France, Iceland, New Zealand, Portugal,[1] Russia, Italy, Slovak Republic
488	Russia	Portugal,[1] Australia, Italy, Slovak Republic, Luxembourg, Spain, Lithuania, Hungary
487	Italy	Portugal,[1] Australia, Russia, Slovak Republic, Luxembourg, Spain, Lithuania, Hungary, United States[1]
486	Slovak Republic	Portugal,[1] Australia, Russia, Italy, Luxembourg, Spain, Lithuania, Hungary, United States[1]
483	Luxembourg	Russia, Italy, Slovak Republic, Spain, Lithuania, Hungary, United States[1]
481	Spain	Russia, Italy, Slovak Republic, Luxembourg, Lithuania, Hungary, United States[1]
481	Lithuania	Russia, Italy, Slovak Republic, Luxembourg, Spain, Hungary, United States[1]
481	Hungary	Russia, Italy, Slovak Republic, Luxembourg, Spain, Lithuania, United States[1]
478	United States[1]	Italy, Slovak Republic, Luxembourg, Spain, Lithuania, Hungary, Belarus, Malta
472	Belarus	United States,[1] Malta
472	Malta	United States,[1] Belarus
464	Croatia	Israel
463	Israel	Croatia
454	Turkey	Ukraine, Greece, Cyprus, Serbia
453	Ukraine	Turkey, Greece, Cyprus, Serbia
451	Greece	Turkey, Ukraine, Cyprus, Serbia
451	Cyprus	Turkey, Ukraine, Greece, Serbia
448	Serbia	Turkey, Ukraine, Greece, Cyprus, Malaysia
440	Malaysia	Serbia, Albania, Bulgaria, United Arab Emirates, Romania
437	Albania	Malaysia, Bulgaria, United Arab Emirates, Romania
436	Bulgaria	Malaysia, Albania, United Arab Emirates, Brunei Darussalam, Romania, Montenegro
435	United Arab Emirates	Malaysia, Albania, Bulgaria, Romania
430	Brunei Darussalam	Bulgaria, Romania, Montenegro
430	Romania	Malaysia, Albania, Bulgaria, United Arab Emirates, Brunei Darussalam, Montenegro, Kazakhstan, Moldova, Baku (Azerbaijan), Thailand
430	Montenegro	Bulgaria, Brunei Darussalam, Romania

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.5.

StatLink https://doi.org/10.1787/888934028254

• • •

Table I.4.2 [2/2] **Comparing countries' and economies' performance in mathematics**

(blue)	Statistically significantly **above** the OECD average
(white)	**Not statistically significantly different** from the OECD average
(grey)	Statistically significantly **below** the OECD average

Mean score	Comparison country/economy	Countries and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
423	Kazakhstan	Romania, Moldova, Baku (Azerbaijan), Thailand, Uruguay, Chile
421	Moldova	Romania, Kazakhstan, Baku (Azerbaijan), Thailand, Uruguay, Chile
420	Baku (Azerbaijan)	Romania, Kazakhstan, Moldova, Thailand, Uruguay, Chile, Qatar
419	Thailand	Romania, Kazakhstan, Moldova, Baku (Azerbaijan), Uruguay, Chile, Qatar
418	Uruguay	Kazakhstan, Moldova, Baku (Azerbaijan), Thailand, Chile, Qatar
417	Chile	Kazakhstan, Moldova, Baku (Azerbaijan), Thailand, Uruguay, Qatar
414	Qatar	Baku (Azerbaijan), Thailand, Uruguay, Chile, Mexico
409	Mexico	Qatar, Bosnia and Herzegovina, Costa Rica
406	Bosnia and Herzegovina	Mexico, Costa Rica, Peru, Jordan
402	Costa Rica	Mexico, Bosnia and Herzegovina, Peru, Jordan, Georgia, Lebanon
400	Peru	Bosnia and Herzegovina, Costa Rica, Jordan, Georgia, North Macedonia, Lebanon
400	Jordan	Bosnia and Herzegovina, Costa Rica, Peru, Georgia, North Macedonia, Lebanon
398	Georgia	Costa Rica, Peru, Jordan, North Macedonia, Lebanon, Colombia
394	North Macedonia	Peru, Jordan, Georgia, Lebanon, Colombia
393	Lebanon	Costa Rica, Peru, Jordan, Georgia, North Macedonia, Colombia
391	Colombia	Georgia, North Macedonia, Lebanon
384	Brazil	Argentina, Indonesia
379	Argentina	Brazil, Indonesia, Saudi Arabia
379	Indonesia	Brazil, Argentina, Saudi Arabia
373	Saudi Arabia	Argentina, Indonesia, Morocco
368	Morocco	Saudi Arabia, Kosovo
366	Kosovo	Morocco
353	Panama	Philippines
353	Philippines	Panama
325	Dominican Republic	

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.5.

StatLink https://doi.org/10.1787/888934028254

VARIATION IN PERFORMANCE WITHIN COUNTRIES AND ECONOMIES

While differences in average performance across countries and economies are large, the gap that separates the highest-performing and lowest-performing students within any country is, typically, even larger. In reading, for example, the difference between the 95th percentile of performance (the score above which only 5% of students scored) and the 5th percentile of performance (the score below which only 5% of students scored) was more than 220 score points in all countries and economies; on average across OECD countries, 327 score points separated these extremes (Table I.B1.4). This difference corresponds, typically, to capacities that students develop over the equivalent of several years and grades.[4]

The largest differences between top-performing and low-achieving students were found in Israel, Lebanon, Malta and the United Arab Emirates, meaning that learning outcomes at age 15 in these countries are highly unequal (Table I.B1.4).

The smallest differences between high- and low-achieving students were, typically, found amongst countries and economies with the lowest mean scores. In Kosovo, Morocco and the Philippines, even the highest-performing students scored only around the OECD average. In these countries/economies, the 95th percentile of the reading distribution was close to the average score across OECD countries.

The standard deviation summarises the variation in performance amongst 15-year-old students within each country/economy across the entire distribution. The average standard deviation in reading performance within OECD countries was 99 score points. If the between-country variation was also considered ("OECD total"), the standard deviation across all students in OECD countries was 105 score points. By this measure, the smallest variation in reading proficiency was found in Kosovo (68 score points); several other countries and economies whose mean performance was below the OECD average also have small variations in performance (Figure I.4.1). Amongst high-performing systems, B-S-J-Z (China) (87 score points) stood out for its relatively small variation in performance. This indicates that, more than in other high-performing systems, student performance in B-S-J-Z (China) is consistently high: there are smaller-than-average inequalities in learning outcomes.

Table I.4.3 [1/2] Comparing countries' and economies' performance in science

(blue)	Statistically significantly **above** the OECD average
(white)	**Not statistically significantly different** from the OECD average
(grey)	Statistically significantly **below** the OECD average

Mean score	Comparison country/economy	Countries and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
590	B-S-J-Z (China)	
551	Singapore	
544	Macao (China)	
530	Estonia	Japan
529	Japan	Estonia
522	Finland	Korea, Canada, Hong Kong (China),[1] Chinese Taipei
519	Korea	Finland, Canada, Hong Kong (China),[1] Chinese Taipei
518	Canada	Finland, Korea, Hong Kong (China),[1] Chinese Taipei
517	Hong Kong (China)[1]	Finland, Korea, Canada, Chinese Taipei, Poland
516	Chinese Taipei	Finland, Korea, Canada, Hong Kong (China),[1] Poland
511	Poland	Hong Kong (China),[1] Chinese Taipei, New Zealand, Slovenia, United Kingdom
508	New Zealand	Poland, Slovenia, United Kingdom, Netherlands,[1] Germany, United States[1]
507	Slovenia	Poland, New Zealand, United Kingdom, Netherlands,[1] Germany, Australia, United States[1]
505	United Kingdom	Poland, New Zealand, Slovenia, Netherlands,[1] Germany, Australia, United States,[1] Sweden, Belgium
503	Netherlands[1]	New Zealand, Slovenia, United Kingdom, Germany, Australia, United States,[1] Sweden, Belgium, Czech Republic
503	Germany	New Zealand, Slovenia, United Kingdom, Netherlands,[1] Australia, United States,[1] Sweden, Belgium, Czech Republic, Ireland, Switzerland
503	Australia	Slovenia, United Kingdom, Netherlands,[1] Germany, United States,[1] Sweden, Belgium
502	United States[1]	New Zealand, Slovenia, United Kingdom, Netherlands,[1] Germany, Australia, Sweden, Belgium, Czech Republic, Ireland, Switzerland
499	Sweden	United Kingdom, Netherlands,[1] Germany, Australia, United States,[1] Belgium, Czech Republic, Ireland, Switzerland, France, Denmark, Portugal[1]
499	Belgium	United Kingdom, Netherlands,[1] Germany, Australia, United States,[1] Sweden, Czech Republic, Ireland, Switzerland, France
497	Czech Republic	Netherlands,[1] Germany, United States,[1] Sweden, Belgium, Ireland, Switzerland, France, Denmark, Portugal,[1] Norway, Austria
496	Ireland	Germany, United States,[1] Sweden, Belgium, Czech Republic, Switzerland, France, Denmark, Portugal,[1] Norway, Austria
495	Switzerland	Germany, United States,[1] Sweden, Belgium, Czech Republic, Ireland, France, Denmark, Portugal,[1] Norway, Austria
493	France	Sweden, Belgium, Czech Republic, Ireland, Switzerland, Denmark, Portugal,[1] Norway, Austria
493	Denmark	Sweden, Czech Republic, Ireland, Switzerland, France, Portugal,[1] Norway, Austria
492	Portugal[1]	Sweden, Czech Republic, Ireland, Switzerland, France, Denmark, Norway, Austria, Latvia
490	Norway	Czech Republic, Ireland, Switzerland, France, Denmark, Portugal,[1] Austria, Latvia
490	Austria	Czech Republic, Ireland, Switzerland, France, Denmark, Portugal,[1] Norway, Latvia
487	Latvia	Portugal,[1] Norway, Austria, Spain
483	Spain	Latvia, Lithuania, Hungary, Russia
482	Lithuania	Spain, Hungary, Russia
481	Hungary	Spain, Lithuania, Russia, Luxembourg
478	Russia	Spain, Lithuania, Hungary, Luxembourg, Iceland, Croatia, Belarus
477	Luxembourg	Hungary, Russia, Iceland, Croatia
475	Iceland	Russia, Luxembourg, Croatia, Belarus, Ukraine
472	Croatia	Russia, Luxembourg, Iceland, Belarus, Ukraine, Turkey, Italy
471	Belarus	Russia, Iceland, Croatia, Ukraine, Turkey, Italy
469	Ukraine	Iceland, Croatia, Belarus, Turkey, Italy, Slovak Republic, Israel
468	Turkey	Croatia, Belarus, Ukraine, Italy, Slovak Republic, Israel
468	Italy	Croatia, Belarus, Ukraine, Turkey, Slovak Republic, Israel
464	Slovak Republic	Ukraine, Turkey, Italy, Israel
462	Israel	Ukraine, Turkey, Italy, Slovak Republic, Malta
457	Malta	Israel, Greece
452	Greece	Malta
444	Chile	Serbia, Cyprus, Malaysia
440	Serbia	Chile, Cyprus, Malaysia, United Arab Emirates
439	Cyprus	Chile, Serbia, Malaysia
438	Malaysia	Chile, Serbia, Cyprus, United Arab Emirates
434	United Arab Emirates	Serbia, Malaysia, Brunei Darussalam, Jordan, Moldova, Romania
431	Brunei Darussalam	United Arab Emirates, Jordan, Moldova, Thailand, Uruguay, Romania, Bulgaria
429	Jordan	United Arab Emirates, Brunei Darussalam, Moldova, Thailand, Uruguay, Romania, Bulgaria
428	Moldova	United Arab Emirates, Brunei Darussalam, Jordan, Thailand, Uruguay, Romania, Bulgaria
426	Thailand	Brunei Darussalam, Jordan, Moldova, Uruguay, Romania, Bulgaria, Mexico
426	Uruguay	Brunei Darussalam, Jordan, Moldova, Thailand, Romania, Bulgaria, Mexico
426	Romania	United Arab Emirates, Brunei Darussalam, Jordan, Moldova, Thailand, Uruguay, Bulgaria, Mexico, Qatar, Albania, Costa Rica
424	Bulgaria	Brunei Darussalam, Jordan, Moldova, Thailand, Uruguay, Romania, Mexico, Qatar, Albania, Costa Rica
419	Mexico	Thailand, Uruguay, Romania, Bulgaria, Qatar, Albania, Costa Rica, Montenegro, Colombia
419	Qatar	Romania, Bulgaria, Mexico, Albania, Costa Rica, Colombia
417	Albania	Romania, Bulgaria, Mexico, Qatar, Costa Rica, Montenegro, Colombia, North Macedonia
416	Costa Rica	Romania, Bulgaria, Mexico, Qatar, Albania, Montenegro, Colombia, North Macedonia

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.6.

StatLink https://doi.org/10.1787/888934028273

• • •

Table I.4.3 [2/2] Comparing countries' and economies' performance in science

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

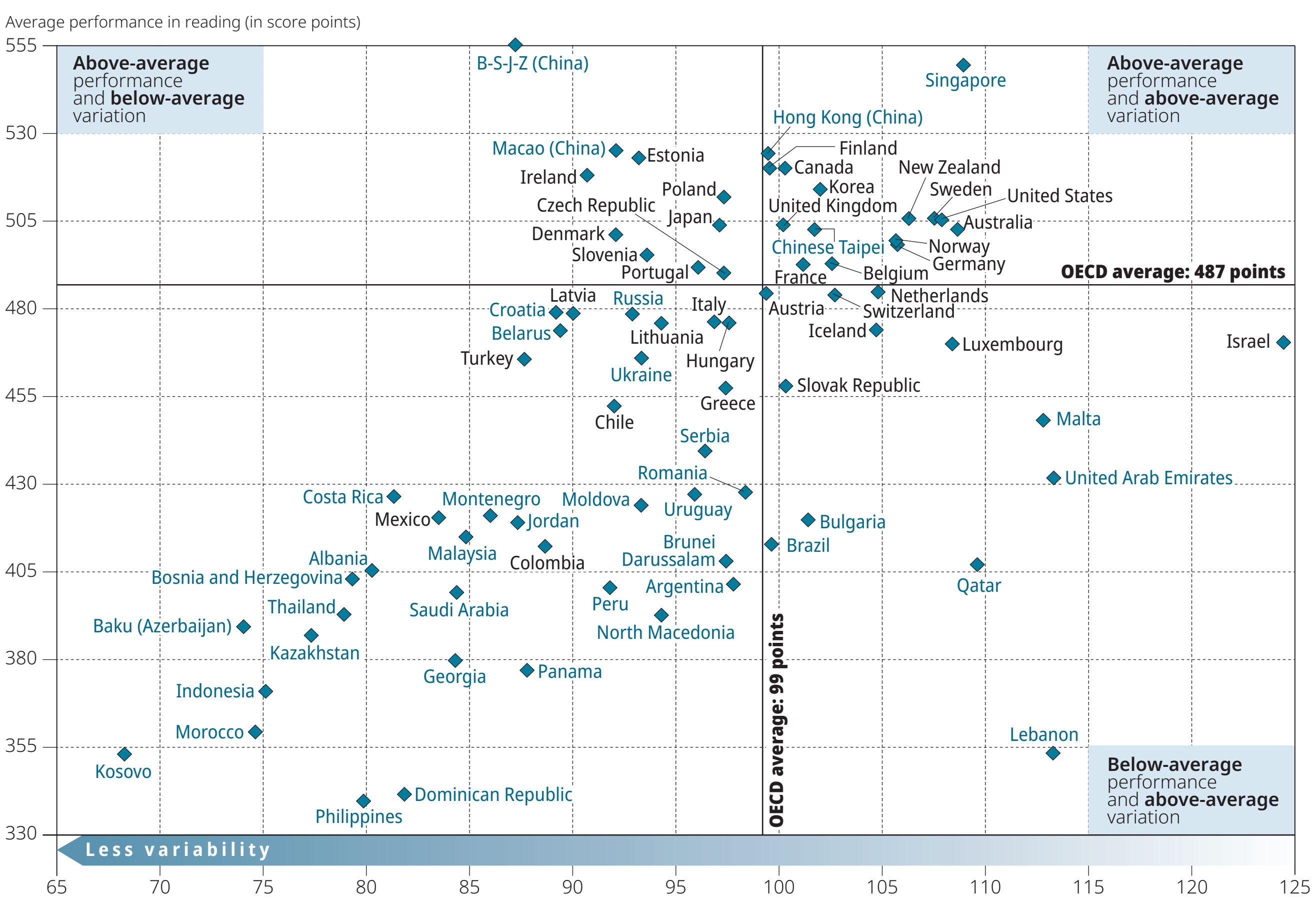
Mean score	Comparison country/economy	Countries and economies whose mean score is not statistically significantly different from the comparison country's/economy's score
415	Montenegro	Mexico, Albania, Costa Rica, Colombia, North Macedonia
413	Colombia	Mexico, Qatar, Albania, Costa Rica, Montenegro, North Macedonia
413	North Macedonia	Albania, Costa Rica, Montenegro, Colombia
404	Peru	Argentina, Brazil, Bosnia and Herzegovina, Baku (Azerbaijan)
404	Argentina	Peru, Brazil, Bosnia and Herzegovina, Baku (Azerbaijan)
404	Brazil	Peru, Argentina, Bosnia and Herzegovina, Baku (Azerbaijan)
398	Bosnia and Herzegovina	Peru, Argentina, Brazil, Baku (Azerbaijan), Kazakhstan, Indonesia
398	Baku (Azerbaijan)	Peru, Argentina, Brazil, Bosnia and Herzegovina, Kazakhstan, Indonesia
397	Kazakhstan	Bosnia and Herzegovina, Baku (Azerbaijan), Indonesia
396	Indonesia	Bosnia and Herzegovina, Baku (Azerbaijan), Kazakhstan
386	Saudi Arabia	Lebanon, Georgia
384	Lebanon	Saudi Arabia, Georgia, Morocco
383	Georgia	Saudi Arabia, Lebanon, Morocco
377	Morocco	Lebanon, Georgia
365	Kosovo	Panama
365	Panama	Kosovo, Philippines
357	Philippines	Panama
336	Dominican Republic	

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Source: OECD, PISA 2018 Database, Table I.B1.6.

StatLink https://doi.org/10.1787/888934028273

Figure I.4.1 Average performance in reading and variation in performance



Source: OECD, PISA 2018 Database, Table I.B1.4.

StatLink https://doi.org/10.1787/888934028349

In contrast, Singapore, with mean performance similar to that of B-S-J-Z (China), had one of the widest variations in reading performance (109 score points; the variation in mathematics and in science was closer to the OECD average). This large variation in reading performance in Singapore may be related to the diversity of students' linguistic backgrounds. As shown at the end of this chapter, 43% of students in Singapore reported that they do not speak the test language at home (Figure I.4.11).[5] (Demographic and socio-economic factors related to variations in performance within countries/economies are more extensively analysed in *PISA 2018 Results [Volume II]: Where All Students Can Succeed* [OECD, 2019[1]]).

RANKING COUNTRIES' AND ECONOMIES' PERFORMANCE IN PISA

The goal of PISA is to provide useful information to educators and policy makers concerning the strengths and weaknesses of their country's education system, the progress made over time, and opportunities for improvement. When ranking countries, economies and education systems in PISA, it is important to consider the social and economic context in which education takes place. Moreover, many countries and economies score at similar levels; small differences that are not statistically significant or practically meaningful should not be overly emphasised.

Table I.4.4, Table I.4.5 and Table I.4.6 show, for each country and economy, an estimate of where its mean performance ranks amongst all other countries and economies that participate in PISA as well as, for OECD countries, amongst all OECD countries. Because mean-score estimates are derived from samples and are thus associated with statistical uncertainty, it is often not possible to determine an exact ranking for all countries and economies. However, it is possible to identify the range of possible rankings for the country's/economy's mean performance.[6] This range of ranks can be wide, particularly for countries/economies whose mean scores are similar to those of many other countries/economies.[7]

Table I.4.4, Table I.4.5 and Table I.4.6 also include, for countries where the sampling design supports such reporting, the results of cities, regions, states or other subnational entities within the country.[8] For these subnational entities (whose results are reported in Annex B2), a rank order was not estimated. Still, the mean score and its confidence interval allow for a comparison of performance with that of countries and economies. For example, Alberta (Canada) scored below top-performers B-S-J-Z (China) and Singapore, but close to Macao (China) in reading. These subnational results also highlight differences within countries that are often as large as between-country differences in performance. In reading, for example, more than 40 score points separated the mean performance of Alberta and the mean performance of New Brunswick in Canada, and even larger differences were observed between Astana and the Atyrau region of Kazakhstan.

A CONTEXT FOR COUNTRIES' PERFORMANCE IN PISA

Comparing the performance of students across vastly diverse countries poses numerous challenges. In any classroom, students with varying abilities, attitudes and social backgrounds are required to respond to the same set of tasks when sitting a test. When comparing the performance of schools in an education system, the same test is used across schools that may differ significantly in the structure and sequencing of their curriculum, in their pedagogical emphasis, in the instructional methods applied, and in the demographic and social contexts of their student population. Comparing the performance of education systems across countries adds further layers of complexity because students are given tests in different languages, and because the social, economic and cultural context of the countries that are being compared are often very different.

However, while students within a country may learn in different contexts according to their home environment and the school they attend, their performance is measured against common standards. And when they become adults, they will all face common challenges and will often have to compete for the same jobs. Similarly, in a global society and economy, the success of education systems in preparing students for life is no longer measured against locally established benchmarks, but increasingly against benchmarks that are common to all education systems around the world. As difficult as international comparisons are, comparisons with the best-performing systems provide important information for educators, and PISA goes to considerable lengths to ensure that such comparisons are valid and fair (see also Annex A6).

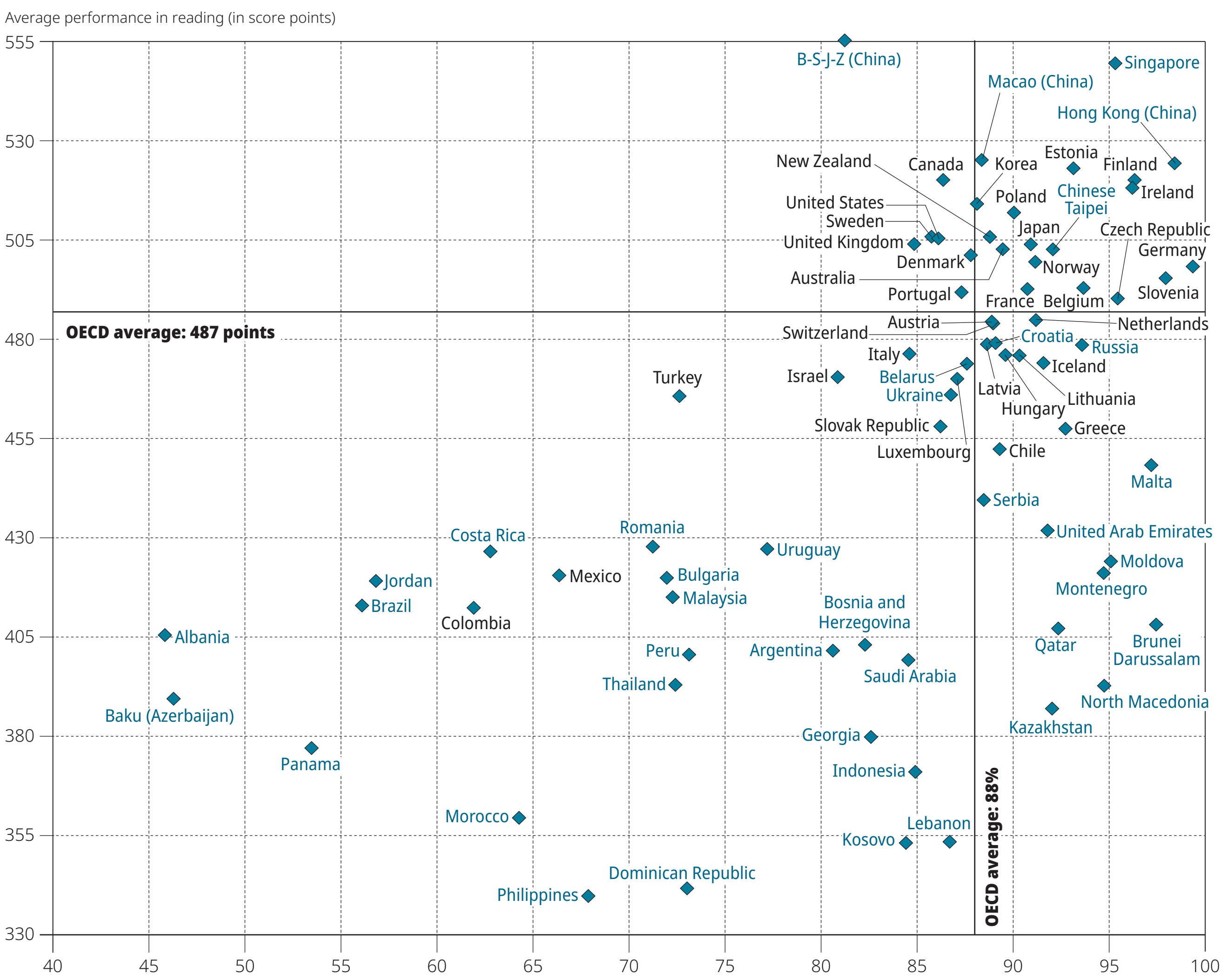
This section discusses countries' mean reading performance in PISA in the context of important economic, demographic and social factors that can influence the assessment results (results are similar for mathematics and science). It provides a context for interpreting the results that are presented above and in the following chapters.

PISA's stringent sampling standards limit the possible exclusion of students and schools and the impact of non-response. These standards are applied to ensure that the results support conclusions that are valid for the PISA target population when comparing adjudicated countries, economies and subnational entities. Chapter 3 provides a definition of the PISA target population, which is the relevant population when comparing school systems.

But when interpreting PISA results with regard to the overall population of 15-year-olds, sample coverage must be assessed with respect to this wider population. Coverage Index 3, discussed in Chapter 3, provides an estimate of the share of the 15-year-old age cohort covered by PISA. In 2018, it varied from 46% in Baku (Azerbaijan) and 53% in Panama to close to 100% in Germany, Hong Kong (China) and Slovenia. While the PISA results are representative of the target population in all adjudicated countries/economies, they cannot be readily generalised to the entire population of 15-year-olds in countries where many young people of that age are not enrolled in lower or upper secondary school. The mean scores of 15-year-old students in countries with a low Coverage Index 3 are typically below average (Figure I.4.2); but the mean scores amongst all 15-year-olds may be even lower if the reading, mathematics and science competences of the 15-year-olds who were not eligible to sit the PISA test were, on average, below those of eligible 15-year-olds.[9] The following chapters (Chapters 5 through 10) discuss several ways of accounting for the share of 15-year-olds who were not covered by the PISA sample when comparing results across countries and over time.

Variations in population coverage are not the only differences that must be borne in mind when comparing results across countries. As discussed in *PISA 2018 Results (Volume II): Where All Students Can Succeed* (OECD, 2019[1]), a family's wealth is related to its children's performance in school, but the strength of this relationship varies markedly across countries. Similarly, the relative prosperity of some countries allows them to spend more on education, while other countries find themselves constrained by a lower national income. It is therefore important to keep the national income of countries in mind when interpreting the performance of middle-income countries, such as Colombia, Moldova, Morocco and the Philippines, compared with high-income countries (defined by the World Bank as countries whose per capita income was above USD 12 375 in 2018).[10]

Figure I.4.2 **Reading performance and coverage of the population of 15-year-olds in the PISA sample**



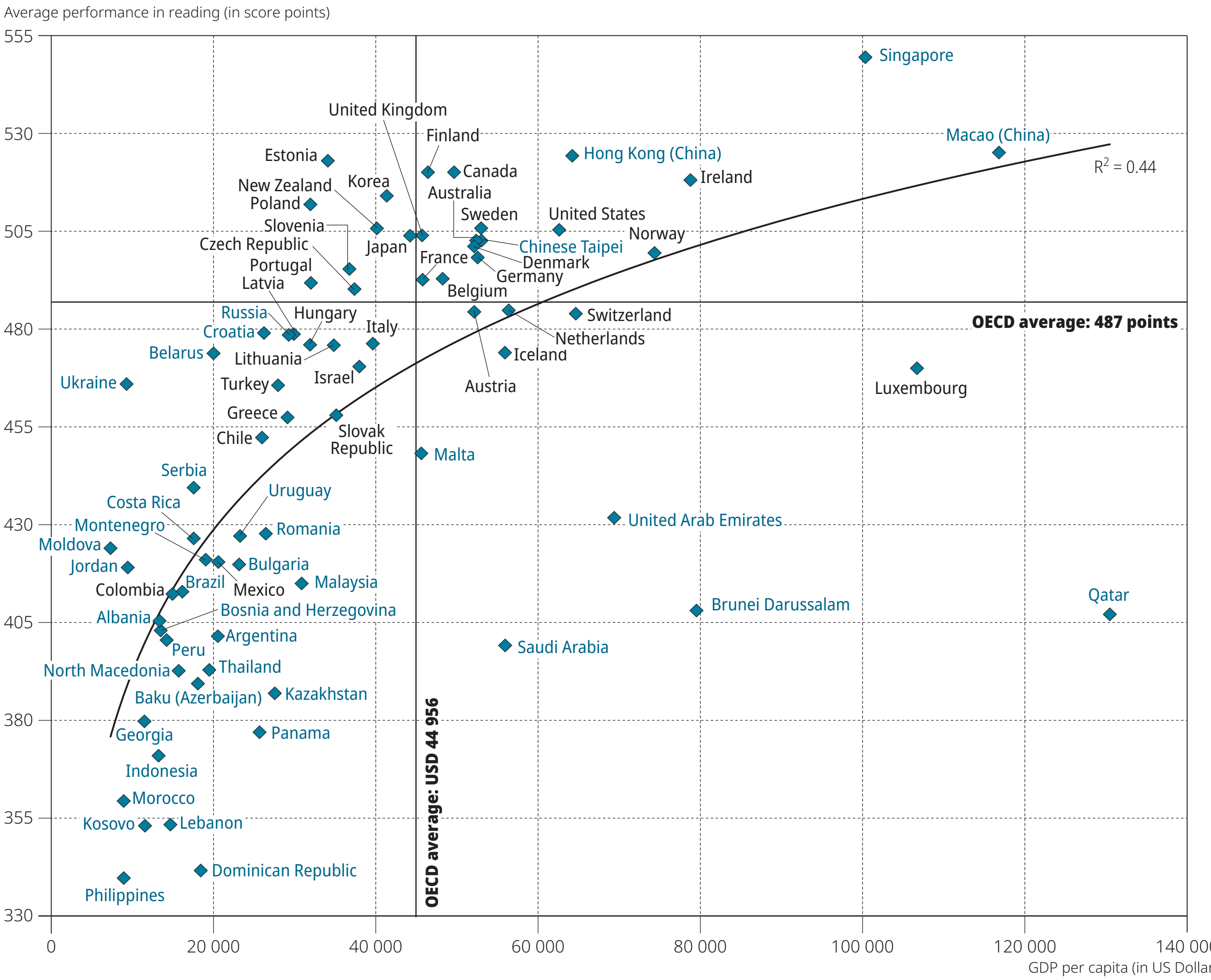
Source: OECD, PISA 2018 Database, Tables I.B1.4 and I.A2.1.

StatLink https://doi.org/10.1787/888934028368

Resources available and invested in education

Figure I.4.3 displays the relationship between national income, as measured by per capita GDP, and students' average reading performance.[11] The figure also shows a trend line that summarises this relationship. The relationship suggests that 44% of the variation in countries'/economies' mean scores is related to per capita GDP (33% in OECD countries). Countries with higher national incomes thus tend to score higher in PISA, even if the chart provides no indications about the causal nature of this relationship. The figure also shows that, although their average performance lies below the OECD average, some countries, including Belarus, Croatia and Ukraine, performed better than other countries at similar levels of economic development.

Figure I.4.3 **Mean reading performance and per capita GDP**



Source: OECD, PISA 2018 Database, Tables I.B1.4 and B3.1.4.

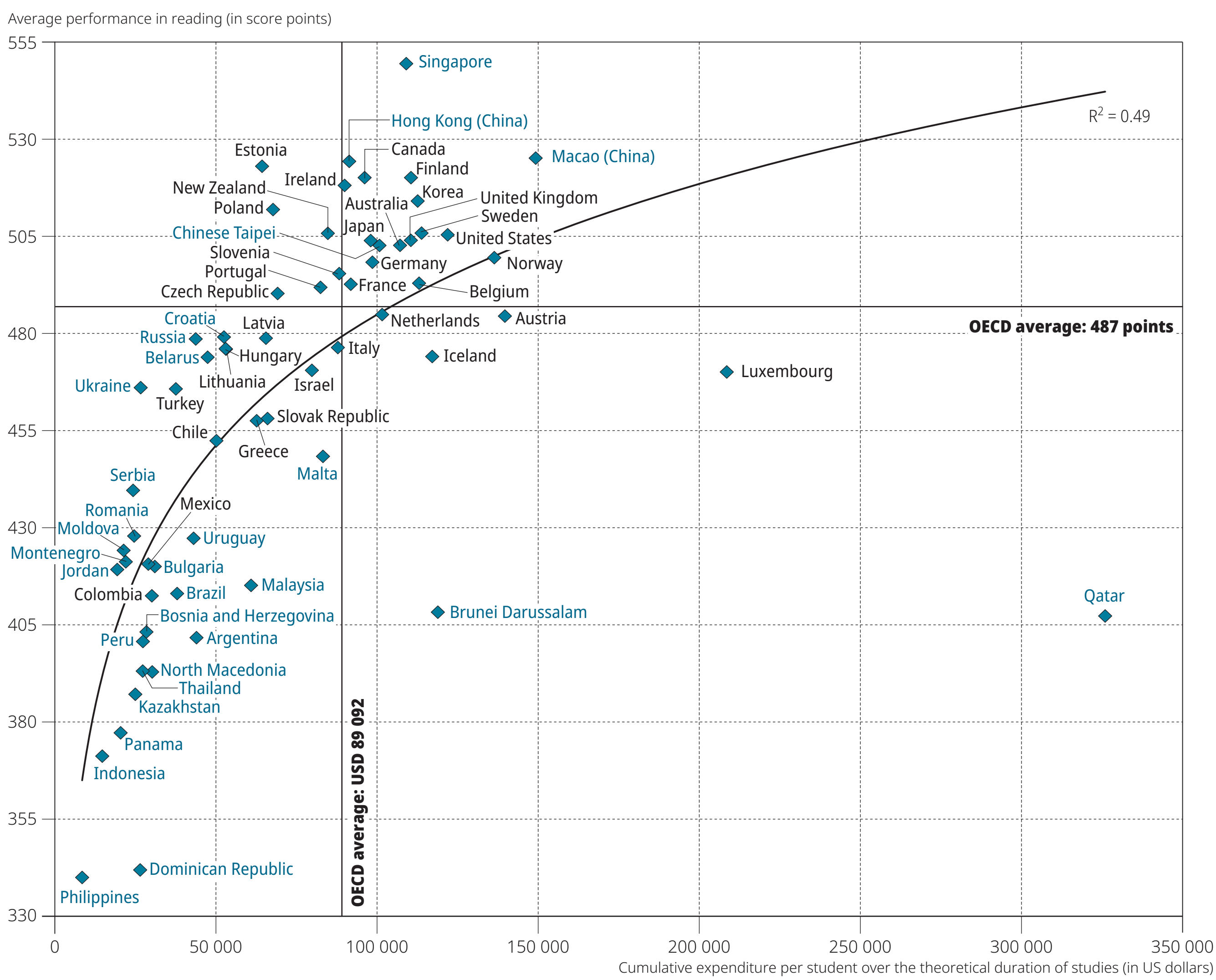
StatLink https://doi.org/10.1787/888934028387

While per capita GDP reflects the potential resources available for education in each country, it does not directly measure the financial resources actually invested in education. Figure I.4.4 compares countries' cumulative spending per student from the age of six up to the age of 15, with average student performance in reading.[12]

Figure I.4.4 shows a positive relationship between spending per student and mean reading performance. As expenditure on educational institutions per student increases, so does a country's mean performance; but the rate of increase diminishes quickly. Expenditure per student accounts for 49% of the variation in mean performance between countries/economies (39% in OECD countries).[13] Relatively low spending per student needs to be taken into account when interpreting the low performance of countries such as Indonesia and the Philippines. But above USD 50 000 per student (after accounting for purchasing power parities [PPP]), a level of cumulative expenditure reached by all OECD countries except Colombia, Mexico and Turkey, spending is much less related to performance. Indeed, Estonia, which spends around USD 64 000 per student (compared to an OECD average expenditure of about USD 89 000), was one of the top-performing OECD countries in reading, mathematics and science in PISA 2018. This shows that, while education needs to be adequately resourced, and is often under-resourced in developing countries, a high level of spending per student is not required to achieve excellence in education.

In most countries, students and their families do not bear the full costs of their primary and secondary education, and often do not pay directly for it, as compulsory education is typically paid for through taxes. But students and their families directly invest their time in education. PISA 2015 highlighted significant differences in the hours of instruction per week among 15-year-old students. Students in Beijing-Shanghai-Jiangsu-Guangdong (China) (hereafter "B-S-J-G [China]"), Chile, Costa Rica, Korea, Chinese Taipei, Thailand and Tunisia spent at least 30 hours per week in regular lessons (all subjects combined), while students in Brazil, Bulgaria, Finland, Lithuania, the Slovak Republic and Uruguay spent less than 25 hours per week. Even larger differences were found in the amount of time that students spent learning outside of regular lessons, i.e. doing homework, taking additional instruction or attending private study. All subjects combined, students in B-S-J-G (China), the Dominican Republic, Qatar, Tunisia and the United Arab Emirates reported that they studied at least 25 hours per week in addition to the required school schedule; in Finland, Germany, Iceland, Japan, the Netherlands, Sweden and Switzerland, they studied less than 15 hours per week outside of school (OECD, 2016, pp. 209-217[4]).

Figure I.4.4 **Reading performance and spending on education**

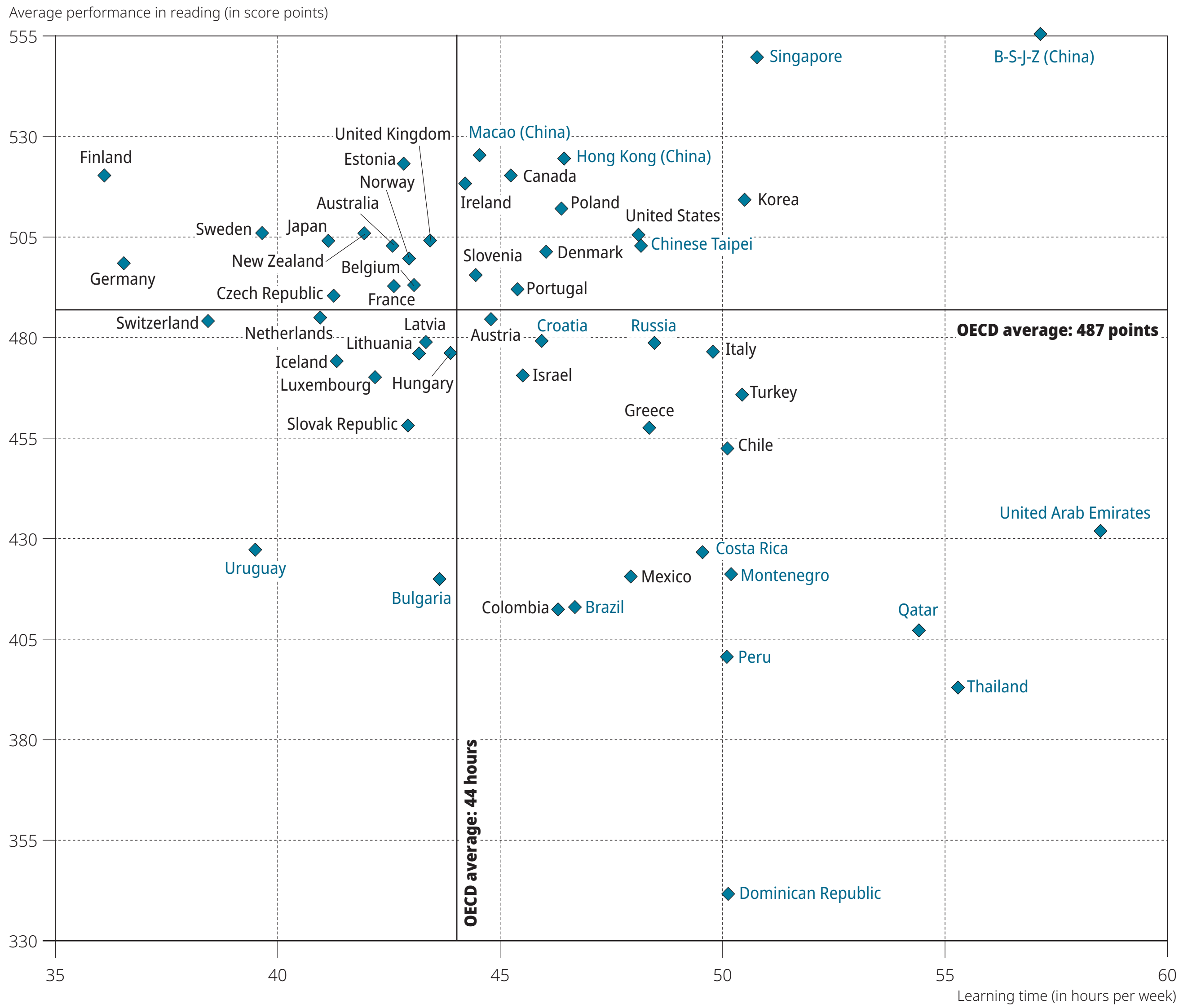


Source: OECD, PISA 2018 Database, Tables I.B1.4 and B3.1.1.

StatLink https://doi.org/10.1787/888934028406

Based on information about learning time collected in PISA 2015,[14] Figure I.4.5 shows the widely varied combinations of total learning time and performance that can be observed across PISA countries and economies. Countries in the upper-left quadrant can be considered more efficient, in that students reach above-average levels of proficiency but devote less time to learning than 15-year-old students on average across OECD countries. This group includes Finland, Germany, Japan and Sweden. By contrast, in several high-performing countries and economies, including B-S-J-Z (China), Korea and Singapore, students reported spending more than 50 hours per week attending regular lessons or in additional learning activities.

Figure I.4.5 **Reading performance and total learning time per week**



Notes: Learning time is based on reports by 15-year-old students in the same country/economy in response to the PISA 2015 questionnaire.

For Beijing-Shanghai-Jiangsu-Zhejiang (China) (labelled as B-J-S-Z [China] on the chart), data on learning time amongst students from Beijing-Shanghai-Jiangsu-Guangdong (China) were used.

Source: OECD, PISA 2018 Database, Table I.B1.4; and OECD, PISA 2015 Database, Figure II.6.23.

StatLink https://doi.org/10.1787/888934028425

The cumulative nature of PISA results

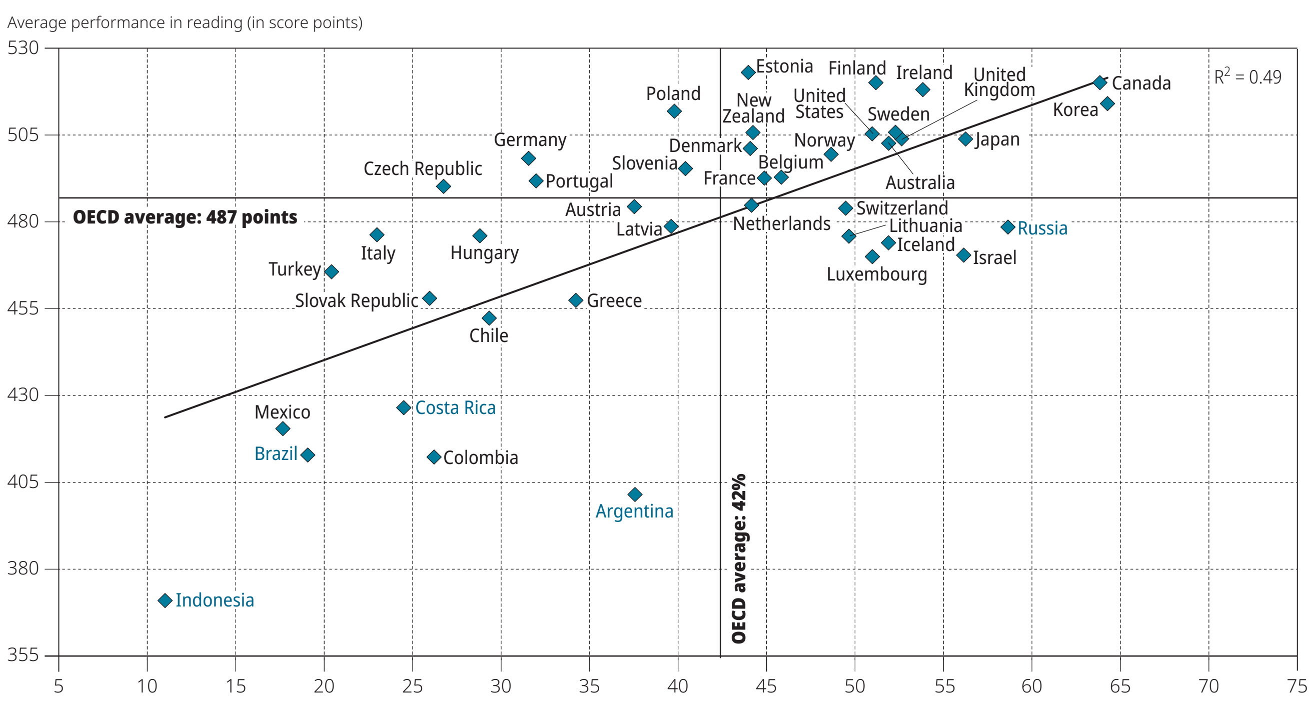
It is not only current economic conditions that matter for education; past economic conditions, and the level of education of previous generations, also influence children's learning outcomes. Indeed, education is a cumulative process: the outcomes of one year of schooling depend on what was learned during the previous year; and the influence of the school environment is compounded by that of the family environment and of the wider social environment in which a child grows up.

There is a close inter-relationship between a student's performance in PISA and his or her parents' level of education (as measured by their educational qualifications); and a similarly close inter-relationship can be expected between countries' performance in PISA and adults' level of education and skills. When it comes to educating their children, countries with more highly educated and skilled adults are at an advantage over countries where parents have less education, or where many adults have low literacy skills. Figure I.4.6 shows the relationship between mean reading performance and the percentage of 35-44 year-olds who have attained tertiary education. This group corresponds roughly to the age group of parents of the 15-year-olds assessed in PISA. According to this simple analysis, the share of tertiary-educated 35-44 year-olds accounts for 49% of the variation between countries/economies (N = 41) in 15-year-old students' mean performance (42% across OECD countries, N = 36). Figure I.4.7 shows the relationship between mean reading performance and the average literacy score of 35-54 year-olds in countries that participated in the Survey of Adult Skills, a product of the OECD Programme for the International Assessment of Adult Competencies (PIAAC).[15] Adult literacy accounts for 58% of the variation in mean performance between countries/economies (N = 35).

When interpreting the performance of 15-year-olds in PISA, it is also important to consider that the results reflect more than the quality of lower secondary schooling (which these students have typically just completed, or are about to complete) or the quality of the upper secondary schools that they may be attending (which, in some cases, they have attended for less than a year). They also reflect the quality of learning in earlier stages of schooling, and the cognitive, emotional and social competences students had acquired before they even entered school.

A clear way of showing this is to compare the mean reading performance of 15-year-olds in PISA with the average reading performance achieved towards the end of primary school by students from a similar birth cohort who participated in the Progress in International Reading Literacy Study (PIRLS) in 2011. Some 42 countries, economies and subnational entities that participated in PISA 2018 also participated in PIRLS 2011, a study developed by the International Association for the Evaluation of Educational Achievement (Mullis et al., 2012[5]). Figure I.4.8 shows a strong correlation between the results of the reading test for 4th-grade students in PIRLS 2011 and the results of the PISA 2018 reading assessment amongst 15-year-old students (variations in PIRLS results can account for about 72% of the variation in PISA reading results across countries and economies). Despite this clear relationship, countries that scored at similar levels in PIRLS – such as the Russian Federation and Singapore, which were amongst the highest-performing countries – can have very different mean scores in PISA. Differences between PISA and PIRLS in countries' relative standing may reflect the influence of the intervening grades on performance, but could also be related to differences in what is measured and in who is assessed.[16]

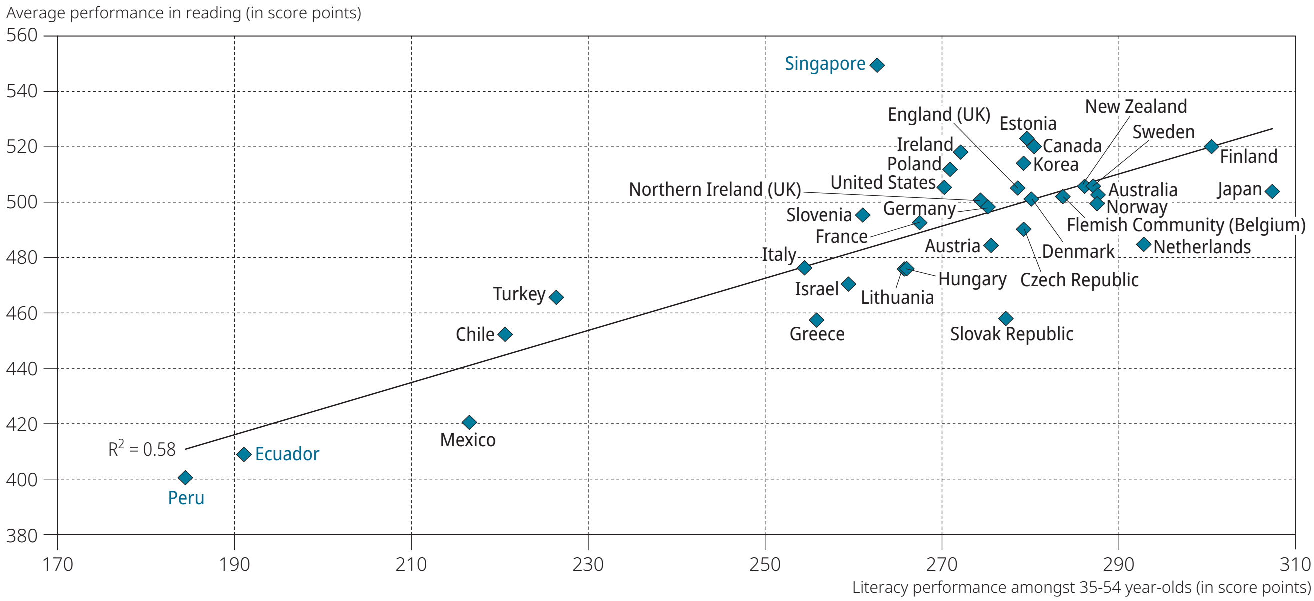
Figure I.4.6 **Reading performance in PISA and educational attainment amongst 35-44 year-olds**



Source: OECD, PISA 2018 Database, Table I.B1.4; OECD (2019[6]), *Education at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/f8d7880d-en.

StatLink https://doi.org/10.1787/888934028444

Figure I.4.7 **Reading performance in PISA and literacy amongst 35-54 year-olds**

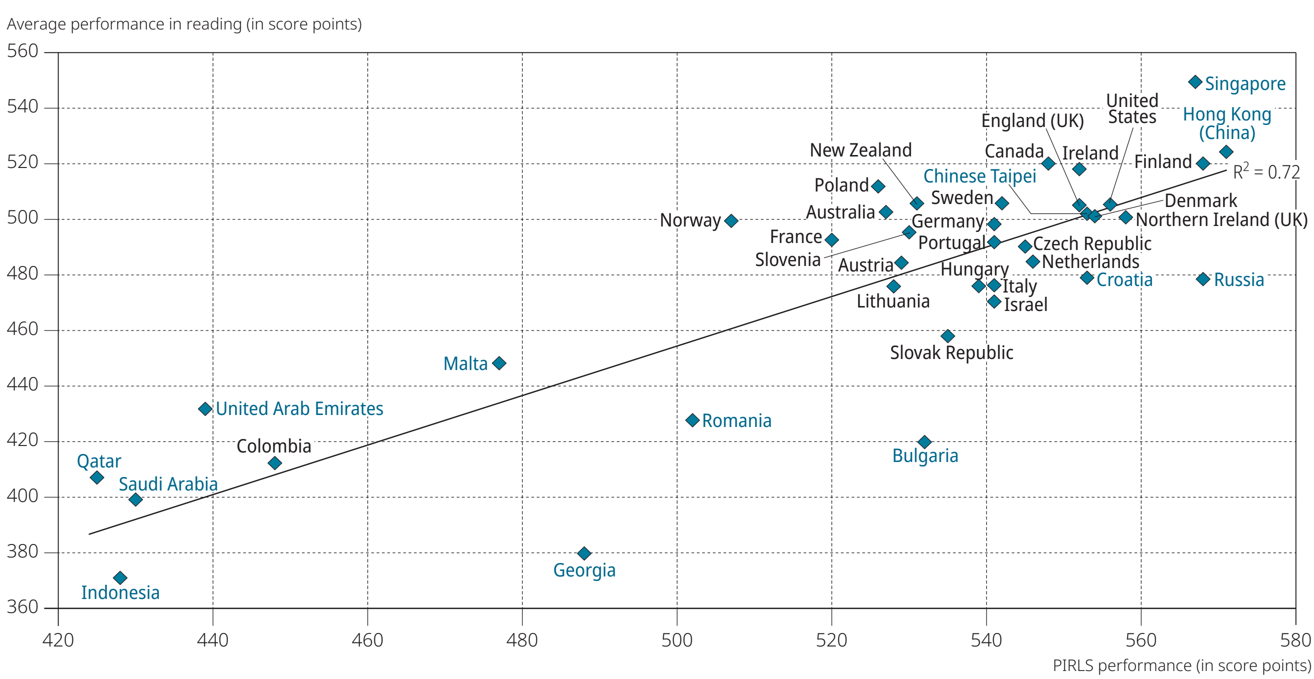


Note: Different countries and regions participated in the Survey of Adult Skills (PIAAC) in different years. In all countries and regions, results for 35-54 year-olds are approximated by the results of adults born between 1964 and 1983. No adjustment was made to account for changes in the skills of these adults, or for changes in the composition of these cohorts, between the year in which the Survey of Adult Skills was conducted and 2018. PISA results for the Flemish community (Belgium) are related to PIAAC results for Flanders (Belgium). PIAAC results for Ecuador are related to the country's results in the PISA for Development assessment (2017). For the United States, PIAAC data refer to 2017.

Source: OECD, PISA 2018 Database, Table I.B1.4; OECD, Survey of Adult Skills (PIAAC) (2011-12, 2014-15, 2017).

StatLink https://doi.org/10.1787/888934028463

Figure I.4.8 **Reading performance in PISA and 4th-graders' performance in PIRLS 2011**



Notes: Only countries and economies with available data are shown.

For Morocco, 6th-grade achievement was used rather than 4th-grade achievement.

Source: OECD, PISA 2018 Database, Table I.B1.4 and Mullis, I. et al. (2012 [5]), *PIRLS 2011 International Results in Reading,* https://timssandpirls.bc.edu/pirls2011/downloads/P11_IR_FullBook.pdf.

StatLink https://doi.org/10.1787/888934028482

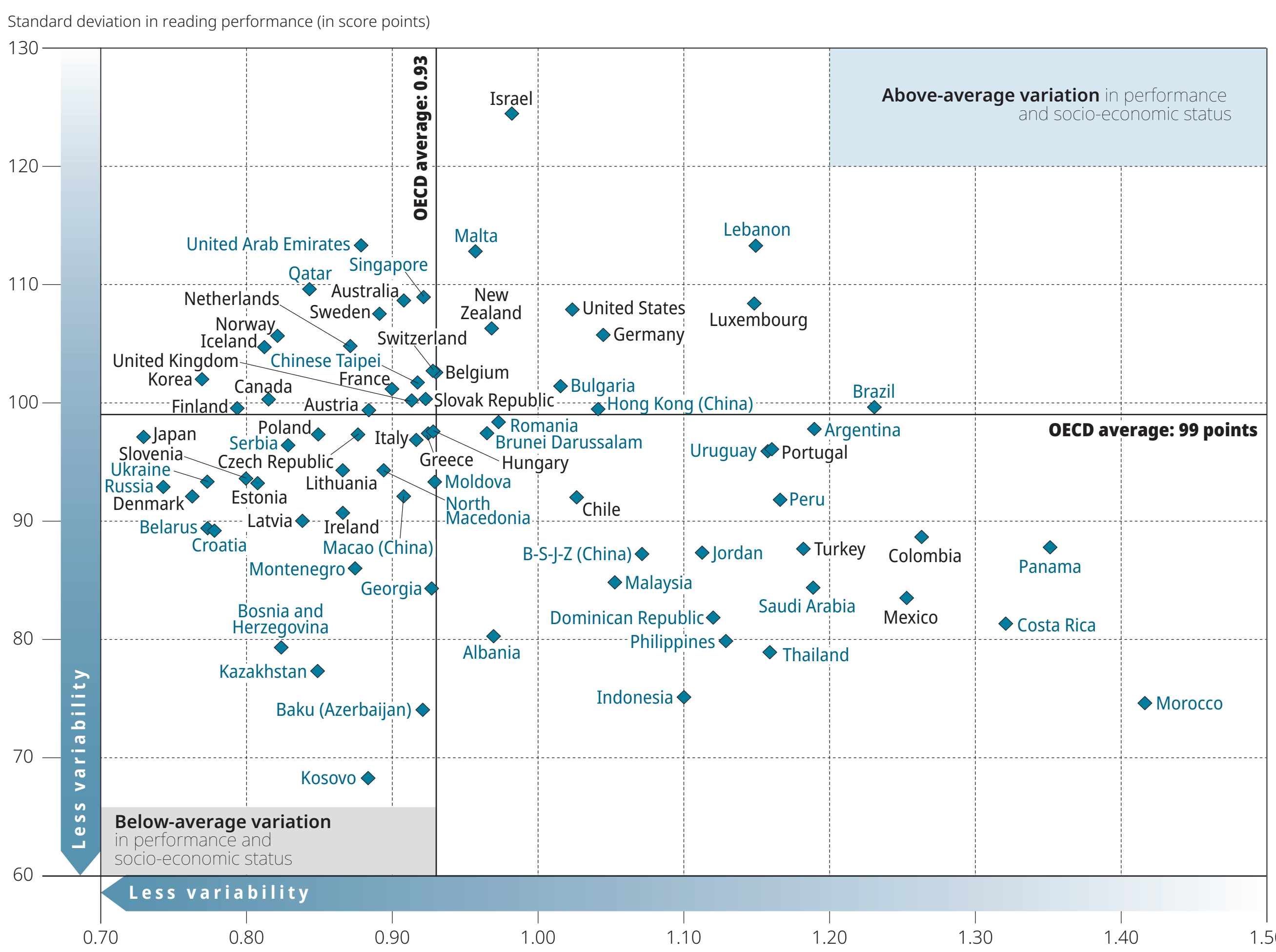
The challenges of student and language diversity

The challenges education systems face cannot be reduced to differences in the overall resources available for schooling or in the extent to which families and society at large support students' acquisition of core skills. Student diversity, related, for example, to socio-economic inequality and students not speaking the language of instruction at home, must also be considered. The challenge for teachers and education systems is to overcome inequalities and at the same time exploit the benefits of diversity in the classroom (OECD, 2010[7]; OECD, 2019[8]).

Figure I.4.9 shows how the standard deviation of reading performance, described earlier, relates to a measure of socio-economic heterogeneity within the country (the standard deviation of the PISA index of economic, social and cultural status); see Chapter 2 in *PISA 2018 Results (Volume II): Where All Students Can Succeed* (OECD, 2019[1]). There is no strong relationship across countries and economies between the magnitude of socio-economic inequalities and the extent to which learning outcomes vary (this also holds after accounting for mean performance in reading). However, some countries (including Brazil, Lebanon and Luxembourg) have comparatively large variations in socio-economic conditions amongst their students, and also larger variations in learning outcomes amongst their students than that observed in countries with similar overall performance or at similar levels of economic development.

How well students read in the language of instruction is influenced by whether they commonly speak that language at home and, more generally, outside of school, and whether specific support is available for bilingual students and for non-native language learners.[17] Specific policies may also be required to help integrate students with an immigrant background into host societies (OECD, 2019[8]); also see *PISA 2018 Results (Volume II): Where All Students Can Succeed* (OECD, 2019[1]), Chapters 9 and 10. But even when such policies are in place, the performance of students who immigrated to the country in which they were assessed can be only partially attributed to their host country's education system.

Figure I.4.9 **Variation in reading performance and in students' socio-economic status**

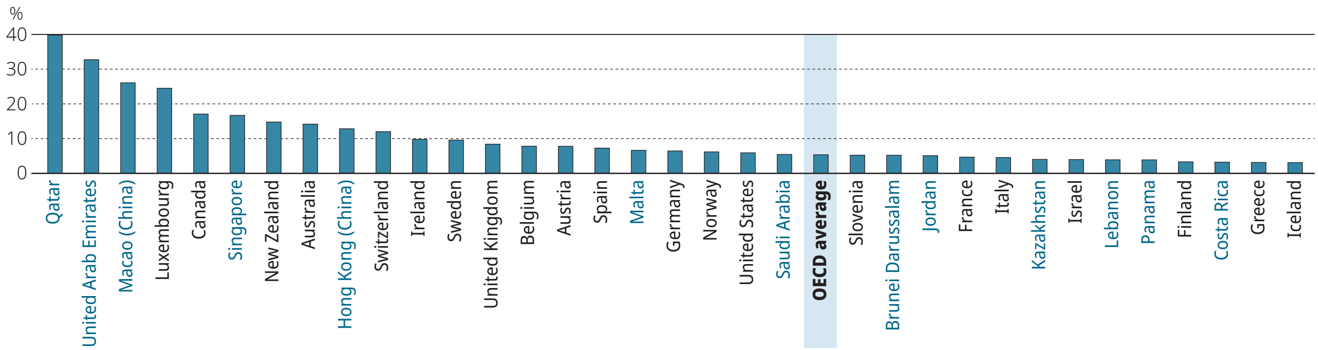


Source: OECD, PISA 2018 Database, Tables I.B1.4 and II.B1.2.1.

StatLink https://doi.org/10.1787/888934028501

Figure I.4.10 **First-generation immigrant students**

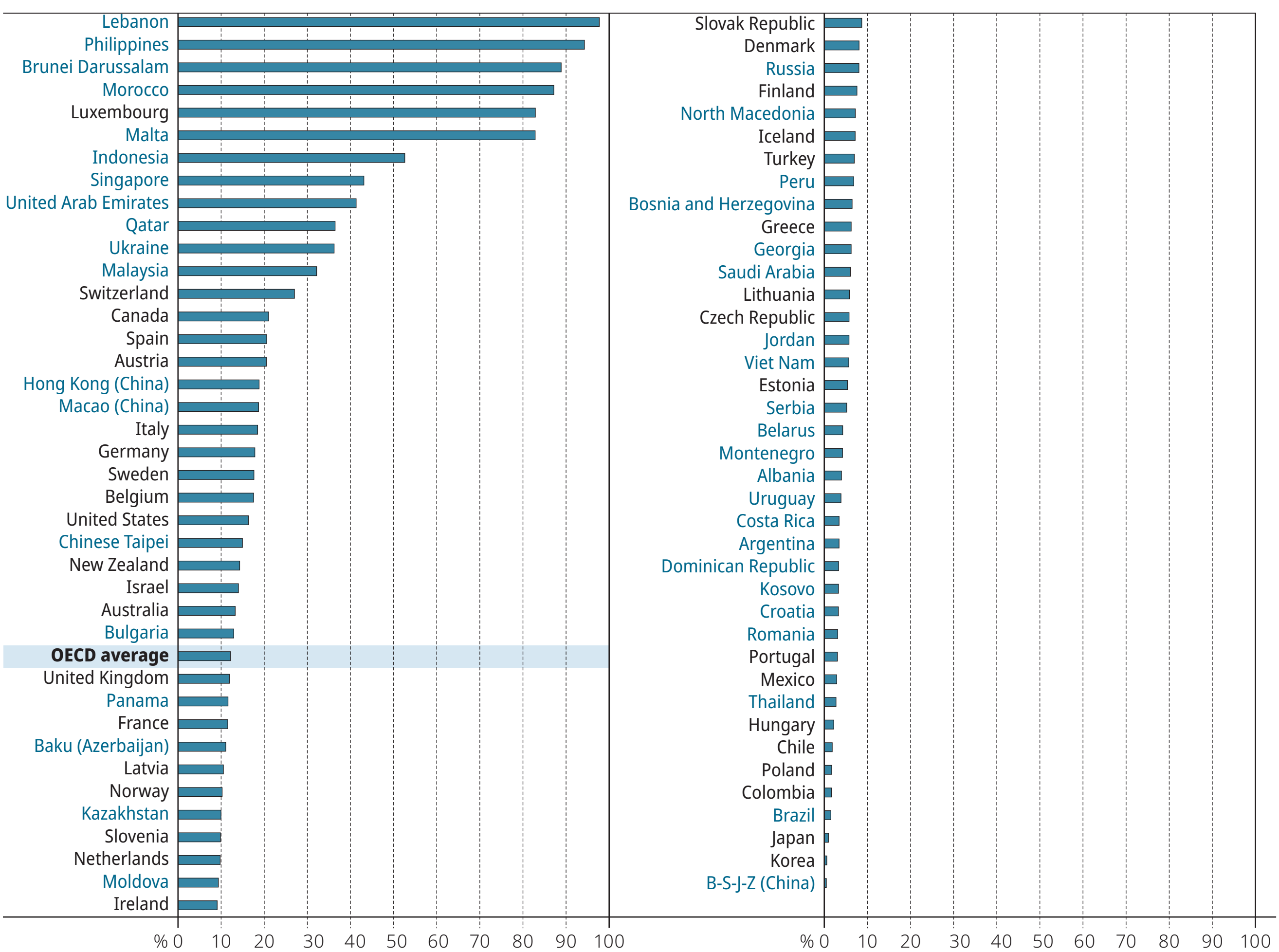
Based on students' reports



Note: Only countries and economies where the percentage of first-generation immigrant students is higher than 3% are shown.
Countries and economies are ranked in descending order of the percentage of first-generation immigrant students.
Source: OECD, PISA 2018 Database, Table II.B1.9.9.
StatLink https://doi.org/10.1787/888934028520

Figure I.4.11 **Students who do not speak the language of instruction at home**

Based on students' reports about what language they speak at home most of the time



Countries and economies are ranked in descending order of the percentage of students who speak, most of the time, a language different from the language of instruction at home.
Source: OECD, PISA 2018 Database, Table II.B1.9.2.
StatLink https://doi.org/10.1787/888934028539

Figure I.4.10 and Figure I.4.11 show the countries where immigration and linguistic diversity are most pronounced.[18] In 2018, more than one in five students in Qatar (40%), the United Arab Emirates (33%), Macao (China) (26%) and Luxembourg (25%) were first-generation immigrants, meaning that they were born outside of the country/economy and their parents were also born outside of the country/economy. In Canada, Singapore, New Zealand, Australia, Hong Kong (China) and Switzerland (in descending order of that share), more than 10% of students were first-generation immigrants. However, some of these immigrants may have already spoken the language of instruction when they arrived. Immigrant students' performance and characteristics are the topic of Chapters 9 and 10 in *PISA 2018 Results (Volume II): Where All Students Can Succeed* (OECD, 2019[1]).

On the other hand, great linguistic diversity may exist even in countries that have relatively small shares of immigrant students. More than 80% of students in Lebanon, the Philippines, Brunei Darussalam, Morocco, Luxembourg and Malta (in descending order of that share), and between 41% and 53% of students in Indonesia, Singapore and the United Arab Emirates reported that, most of the time, they speak a different language at home from the language of instruction.

Table I.4.4 [1/3] Reading performance at national and subnational levels

	Reading scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
B-S-J-Z (China)	555	550 - 561			1	2	1	2
Singapore	549	546 - 553			1	2	1	2
Alberta (Canada)	532	523 - 540						
Macao (China)	525	523 - 528			3	5	3	5
Hong Kong (China)[1]	524	519 - 530			3	7	3	7
Ontario (Canada)	524	517 - 531						
Estonia	523	519 - 527	1	3	3	7	3	7
Canada	520	517 - 524	1	4	4	8	4	8
Finland	520	516 - 525	1	5	4	9	4	9
Québec (Canada)	519	513 - 526						
British Columbia (Canada)	519	511 - 528						
Ireland	518	514 - 522	1	5	5	9	5	9
Nova Scotia (Canada)	516	508 - 523						
Korea	514	508 - 520	2	7	6	11	6	11
Newfoundland and Labrador (Canada)	512	503 - 520						
Poland	512	507 - 517	4	8	8	12	8	12
Sweden	506	500 - 512	6	14	10	19	10	19
New Zealand	506	502 - 510	6	12	10	17	10	17
United States[1]	505	498 - 512	6	15	10	20	10	20
England (United Kingdom)	505	499 - 511						
Scotland (United Kingdom)	504	498 - 510						
United Kingdom	504	499 - 509	7	15	11	20	11	20
Japan	504	499 - 509	7	15	11	20	11	20
Australia	503	499 - 506	8	14	12	19	12	19
Chinese Taipei	503	497 - 508			11	20	11	20
Prince Edward Island (Canada)	503	486 - 519						
Flemish Community (Belgium)	502	495 - 509						
Denmark	501	498 - 505	9	15	13	20	13	20
Northern Ireland (United Kingdom)	501	493 - 509						
Norway	499	495 - 504	10	17	14	22	14	22
Saskatchewan (Canada)	499	493 - 505						
Germany	498	492 - 504	10	19	14	24	14	24
Trento (Italy)	496	491 - 501						
Bolzano (Italy)	495	489 - 502						
Slovenia	495	493 - 498	14	18	19	23	19	23
Manitoba (Canada)	494	488 - 501						
Belgium	493	488 - 497	15	20	20	26	20	26
France	493	488 - 497	15	21	20	26	20	26
Portugal[1]	492	487 - 497	15	21	20	26	20	26
Czech Republic	490	485 - 495	16	22	21	27	21	27
New Brunswick (Canada)	489	482 - 496						
Moscow region (Russia)	486	477 - 495						
Netherlands[1]	485	480 - 490	20	24	24	30	24	30

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028292

...

Table I.4.4 [2/3] Reading performance at national and subnational levels

	Reading scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
Austria	484	479 - 490	20	24	24	30	24	30
Switzerland	484	478 - 490	19	25	24	31	24	31
Wales (United Kingdom)	483	476 - 491						
German-speaking Community (Belgium)	483	474 - 492						
Toscana (Italy)	482	475 - 490						
French Community (Belgium)	481	475 - 487						
Croatia	479	474 - 484			27	36	27	36
Latvia	479	476 - 482	23	27	28	34	28	34
Russia	479	472 - 485			26	36	26	36
Italy	476	472 - 481	23	29	29	37	29	37
Hungary	476	472 - 480	24	29	29	37	29	37
Lithuania	476	473 - 479	24	28	29	36	30	36
Iceland	474	471 - 477	25	29	31	38	31	37
Belarus	474	469 - 479			30	38	30	38
Israel	470	463 - 478	25	31	31	40	31	39
Luxembourg	470	468 - 472	29	31	36	39	36	39
Ukraine	466	459 - 473			36	41		
Turkey	466	461 - 470	30	32	38	41	38	40
Republic of Tatarstan (Russia)	463	456 - 469						
Sardegna (Italy)	462	454 - 470						
Slovak Republic	458	454 - 462	32	34	40	43	40	42
Greece	457	450 - 465	31	34	40	43	39	42
Bogotá (Colombia)	455	444 - 465						
CABA (Argentina)	454	443 - 464						
Chile	452	447 - 457	33	34	42	44	41	43
Malta	448	445 - 452			43	44	42	43
Serbia	439	433 - 446			45	46	44	45
South (Brazil)	432	420 - 444						
United Arab Emirates	432	427 - 436			45	48	44	47
Romania	428	418 - 438			45	55		
Astana (Kazakhstan)	428	413 - 442						
Córdoba (Argentina)	427	418 - 436						
Uruguay	427	422 - 433			46	52	45	49
Costa Rica	426	420 - 433			46	54	45	50
Middle-West (Brazil)	425	407 - 443						
Almaty (Kazakhstan)	424	409 - 440						
Cyprus	424	422 - 427			48	53	46	50
Moldova	424	419 - 429			47	54		
Southeast (Brazil)	424	418 - 430						
Karagandy region (Kazakhstan)	422	409 - 436						
Montenegro	421	419 - 423			50	55	48	51
Mexico	420	415 - 426	35	36	49	57	47	52
Bulgaria	420	412 - 428			48	58	46	53

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue.
Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028292

...

Table I.4.4 [3/3] **Reading performance at national and subnational levels**

	Reading scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
Jordan	419	413 - 425			49	57		
Kostanay region (Kazakhstan)	417	407 - 427						
Malaysia	415	409 - 421			53	58	50	54
DI Yogyakarta (Indonesia)	414	402 - 425						
PBA (Argentina)	413	402 - 424						
Brazil	413	409 - 417			55	59	51	54
North-Kazakhstan region (Kazakhstan)	413	403 - 422						
DKI Jakarta (Indonesia)	412	399 - 426						
Colombia	412	406 - 419	35	36	54	61	51	57
Brunei Darussalam	408	406 - 410			58	61	54	57
Qatar	407	406 - 409			59	62	55	58
Albania	405	402 - 409			59	64	55	59
East-Kazakhstan region (Kazakhstan)	405	392 - 418						
Bosnia and Herzegovina	403	397 - 409			59	65	55	59
Argentina	402	396 - 407			60	66		
Peru	401	395 - 406			61	66	57	60
Saudi Arabia	399	393 - 405			61	66		
Akmola region (Kazakhstan)	395	386 - 404						
Thailand	393	387 - 399			64	69	59	62
North Macedonia	393	391 - 395			66	68		
North (Brazil)	392	379 - 406						
Pavlodar region (Kazakhstan)	391	378 - 403						
Baku (Azerbaijan)	389	384 - 394			66	69	60	62
Northeast (Brazil)	389	381 - 397						
Tucumán (Argentina)	389	379 - 399						
Kazakhstan	387	384 - 390			68	69	61	62
Aktobe region (Kazakhstan)	381	372 - 389						
Georgia	380	376 - 384			70	71	63	64
West-Kazakhstan region (Kazakhstan)	378	369 - 388						
Panama	377	371 - 383			70	72	63	65
Indonesia	371	366 - 376			71	72	64	65
Zhambyl region (Kazakhstan)	369	362 - 376						
South-Kazakhstan region (Kazakhstan)	368	361 - 375						
Kyzyl-Orda region (Kazakhstan)	366	361 - 372						
Mangistau region (Kazakhstan)	361	349 - 372						
Almaty region (Kazakhstan)	360	351 - 369						
Morocco	359	353 - 366			73	74	66	67
Lebanon	353	345 - 362			73	75		
Kosovo	353	351 - 355			74	75	66	67
Atyrau region (Kazakhstan)	344	335 - 352						
Dominican Republic	342	336 - 347			76	77	68	69
Philippines	340	333 - 346			76	77	68	69

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028292

Table I.4.5 [1/3] Mathematics performance at national and subnational levels

	Mathematics scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
B-S-J-Z (China)	591	586 - 596			1	1	1	1
Singapore	569	566 - 572			2	2	2	2
Macao (China)	558	555 - 561			3	4	3	4
Hong Kong (China)[1]	551	545 - 557			3	4	3	4
Québec (Canada)	532	525 - 539						
Chinese Taipei	531	525 - 537			5	7	5	7
Japan	527	522 - 532	1	3	5	8	5	8
Korea	526	520 - 532	1	4	5	9	5	9
Estonia	523	520 - 527	1	4	6	9	6	9
Bolzano (Italy)	521	515 - 528						
Netherlands[1]	519	514 - 524	2	6	7	11	7	11
Trento (Italy)	518	513 - 523						
Flemish Community (Belgium)	518	511 - 524						
Poland	516	511 - 521	4	8	9	13	9	13
Switzerland	515	510 - 521	4	9	9	14	9	14
Ontario (Canada)	513	504 - 521						
Canada	512	507 - 517	5	11	10	16	10	16
Alberta (Canada)	511	501 - 521						
Denmark	509	506 - 513	6	11	11	16	11	16
Slovenia	509	506 - 512	7	11	12	16	12	16
Belgium	508	504 - 513	7	13	12	18	12	18
Finland	507	503 - 511	7	13	12	18	12	18
German-speaking Community (Belgium)	505	495 - 515						
British Columbia (Canada)	504	494 - 515						
England (United Kingdom)	504	498 - 510						
Navarre (Spain)	503	486 - 519						
Castile and León (Spain)	502	493 - 512						
Sweden	502	497 - 508	10	19	15	24	15	24
United Kingdom	502	497 - 507	10	19	15	24	15	24
Norway	501	497 - 505	11	19	16	24	16	24
Germany	500	495 - 505	11	21	16	26	16	26
Ireland	500	495 - 504	12	21	17	26	17	26
Czech Republic	499	495 - 504	12	21	17	26	17	26
Basque Country (Spain)	499	492 - 506						
Austria	499	493 - 505	12	23	17	28	17	28
Cantabria (Spain)	499	484 - 514						
Galicia (Spain)	498	490 - 507						
La Rioja (Spain)	497	478 - 517						
Aragon (Spain)	497	485 - 508						
Latvia	496	492 - 500	15	23	20	28	20	28
Toscana (Italy)	496	487 - 504						
France	495	491 - 500	15	24	20	29	20	29
Iceland	495	491 - 499	16	24	21	29	21	29
French Community (Belgium)	495	490 - 501						
New Zealand	494	491 - 498	18	24	22	29	22	29
Nova Scotia (Canada)	494	482 - 507						
Portugal[1]	492	487 - 498	18	26	23	31	23	31
Northern Ireland (United Kingdom)	492	484 - 500						
Australia	491	488 - 495	20	25	25	31	25	31

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028311

...

Table I.4.5 [2/3] Mathematics performance at national and subnational levels

	Mathematics scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
New Brunswick (Canada)	491	480 - 502						
Asturias (Spain)	491	481 - 500						
Catalonia (Spain)	490	482 - 498						
Scotland (United Kingdom)	489	481 - 497						
Newfoundland and Labrador (Canada)	488	476 - 501						
Russia	488	482 - 494			27	35	27	35
Wales (United Kingdom)	487	479 - 495						
Italy	487	481 - 492	23	29	28	35	28	35
Prince Edward Island (Canada)	487	465 - 508						
Slovak Republic	486	481 - 491	23	29	28	35	28	35
Madrid (Spain)	486	479 - 492						
Saskatchewan (Canada)	485	475 - 495						
Luxembourg	483	481 - 486	25	29	31	36	31	36
Balearic Islands (Spain)	483	472 - 493						
Manitoba (Canada)	482	474 - 489						
Spain	481	479 - 484	26	31	32	37	32	37
Lithuania	481	477 - 485	26	31	32	37	32	37
Hungary	481	477 - 486	26	31	31	37	31	37
Castile-La Mancha (Spain)	479	469 - 489						
United States[1]	478	472 - 485	27	31	32	39	32	39
Murcia (Spain)	474	462 - 485						
Comunidad Valenciana (Spain)	473	465 - 482						
Belarus	472	467 - 477			37	40	37	40
Malta	472	468 - 475			37	39	37	39
Extremadura (Spain)	470	457 - 482						
Andalusia (Spain)	467	459 - 476						
Sardegna (Italy)	467	459 - 475						
Croatia	464	459 - 469			39	41	40	41
Israel	463	456 - 470	32	32	39	42	39	41
Canary Islands (Spain)	460	452 - 469						
Zhambyl region (Kazakhstan)	456	444 - 467						
Turkey	454	449 - 458	33	34	42	46	42	45
Ukraine	453	446 - 460			41	46		
Greece	451	445 - 457	33	34	42	46	42	45
Cyprus	451	448 - 453			42	46	42	45
Astana (Kazakhstan)	450	435 - 466						
Almaty (Kazakhstan)	448	434 - 463						
Serbia	448	442 - 454			42	47	42	46
Kostanay region (Kazakhstan)	448	435 - 461						
Karagandy region (Kazakhstan)	446	431 - 460						
Malaysia	440	435 - 446			46	50	45	49
Pavlodar region (Kazakhstan)	438	426 - 449						
Albania	437	432 - 442			47	51	46	49
East-Kazakhstan region (Kazakhstan)	437	423 - 451						
Bulgaria	436	429 - 444			47	53	46	51
United Arab Emirates	435	431 - 439			47	51	46	50
CABA (Argentina)	434	425 - 444						
North-Kazakhstan region (Kazakhstan)	433	422 - 443						
Melilla (Spain)	432	411 - 452						

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028311

...

Table I.4.5 [3/3] Mathematics performance at national and subnational levels

	Mathematics scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
Brunei Darussalam	430	428 - 432			50	53	49	51
Romania	430	420 - 440			47	56		
DI Yogyakarta (Indonesia)	430	417 - 442						
Montenegro	430	427 - 432			50	53	49	51
Bogotá (Colombia)	430	420 - 439						
Kazakhstan	423	419 - 427			53	57	52	54
DKI Jakarta (Indonesia)	421	406 - 436						
Moldova	421	416 - 425			54	59		
Aktobe region (Kazakhstan)	420	408 - 432						
Baku (Azerbaijan)	420	414 - 425			54	60	52	57
Kyzyl-Orda region (Kazakhstan)	419	403 - 436						
Thailand	419	412 - 425			53	60	52	57
West-Kazakhstan region (Kazakhstan)	418	405 - 430						
Uruguay	418	413 - 423			54	60	52	57
Chile	417	413 - 422	35	35	55	60	53	57
Qatar	414	412 - 417			58	61	55	58
Ceuta (Spain)	411	387 - 435						
Akmola region (Kazakhstan)	411	399 - 424						
Mexico	409	404 - 414	36	36	60	63	57	60
Bosnia and Herzegovina	406	400 - 412			61	65	58	61
Costa Rica	402	396 - 409			61	66	58	62
South-Kazakhstan region (Kazakhstan)	401	390 - 412						
South (Brazil)	401	391 - 412						
Córdoba (Argentina)	400	392 - 409						
Peru	400	395 - 405			62	67	59	62
Jordan	400	393 - 406			62	68		
Almaty region (Kazakhstan)	399	389 - 409						
Georgia	398	392 - 403			63	68	60	63
Middle-West (Brazil)	396	379 - 412						
North Macedonia	394	391 - 398			65	69		
Lebanon	393	386 - 401			63	69		
Southeast (Brazil)	392	386 - 398						
Colombia	391	385 - 397	37	37	66	70	62	64
Mangistau region (Kazakhstan)	391	373 - 409						
PBA (Argentina)	387	377 - 397						
Brazil	384	380 - 388			69	72	64	65
Atyrau region (Kazakhstan)	382	368 - 396						
Argentina	379	374 - 385			70	73		
Indonesia	379	373 - 385			70	73	64	65
Saudi Arabia	373	367 - 379			71	74		
Morocco	368	361 - 374			73	75	66	67
North (Brazil)	366	352 - 380						
Kosovo	366	363 - 369			74	75	66	67
Tucumán (Argentina)	364	354 - 374						
Northeast (Brazil)	363	356 - 371						
Panama	353	348 - 358			76	77	68	69
Philippines	353	346 - 359			76	77	68	69
Dominican Republic	325	320 - 330			78	78	70	70

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028311

Table I.4.6 [1/3] **Science performance at national and subnational levels**

	Science scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
B-S-J-Z (China)	590	585 - 596			1	1	1	1
Singapore	551	548 - 554			2	2	2	2
Macao (China)	544	541 - 546			3	3	3	3
Alberta (Canada)	534	525 - 542						
Estonia	530	526 - 534	1	2	4	5	4	5
Japan	529	524 - 534	1	3	4	6	4	6
Finland	522	517 - 527	2	5	5	9	5	9
Québec (Canada)	522	514 - 529						
Korea	519	514 - 525	3	5	6	10	6	10
Ontario (Canada)	519	511 - 526						
Canada	518	514 - 522	3	5	6	10	6	10
Hong Kong (China)[1]	517	512 - 522			6	11	6	11
British Columbia (Canada)	517	506 - 527						
Chinese Taipei	516	510 - 521			6	11	6	11
Poland	511	506 - 516	5	9	9	14	9	14
Galicia (Spain)	510	503 - 518						
Flemish Community (Belgium)	510	503 - 516						
New Zealand	508	504 - 513	6	10	10	15	10	15
Nova Scotia (Canada)	508	499 - 517						
England (United Kingdom)	507	501 - 513						
Slovenia	507	505 - 509	6	11	11	16	11	16
Newfoundland and Labrador (Canada)	506	494 - 519						
United Kingdom	505	500 - 510	6	14	11	19	11	19
Netherlands[1]	503	498 - 509	7	16	12	21	12	21
Germany	503	497 - 509	7	16	12	21	12	21
Australia	503	499 - 506	8	15	13	20	13	20
United States[1]	502	496 - 509	7	18	12	23	12	23
Prince Edward Island (Canada)	502	484 - 519						
Castile and León (Spain)	501	491 - 511						
Saskatchewan (Canada)	501	493 - 508						
Sweden	499	493 - 505	9	19	14	24	14	24
Belgium	499	494 - 503	11	19	16	24	16	24
Bolzano (Italy)	498	490 - 506						
Czech Republic	497	492 - 502	12	21	17	26	17	26
Asturias (Spain)	496	487 - 505						
Ireland	496	492 - 500	13	21	18	26	18	26
Cantabria (Spain)	495	477 - 513						
Switzerland	495	489 - 501	13	23	18	28	18	28
Trento (Italy)	495	491 - 499						
Aragon (Spain)	493	483 - 504						
France	493	489 - 497	16	23	21	28	21	28
Denmark	493	489 - 496	16	23	21	28	21	28
New Brunswick (Canada)	492	481 - 504						
Navarre (Spain)	492	480 - 504						
Portugal[1]	492	486 - 497	16	24	21	29	21	29
Northern Ireland (United Kingdom)	491	482 - 500						
Norway	490	486 - 495	18	24	23	29	23	29
Scotland (United Kingdom)	490	482 - 498						
Austria	490	484 - 495	18	25	23	30	23	30

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028330

...

Table I.4.6 [2/3] Science performance at national and subnational levels

	Science scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
Manitoba (Canada)	489	482 - 497						
Catalonia (Spain)	489	479 - 498						
Wales (United Kingdom)	488	481 - 496						
Basque Country (Spain)	487	479 - 496						
Latvia	487	484 - 491	21	25	26	30	26	30
Madrid (Spain)	487	481 - 493						
La Rioja (Spain)	487	471 - 502						
French Community (Belgium)	485	479 - 490						
Castile-La Mancha (Spain)	484	473 - 496						
German-speaking Community (Belgium)	483	469 - 498						
Spain	483	480 - 486	24	27	29	32	29	32
Balearic Islands (Spain)	482	472 - 492						
Lithuania	482	479 - 485	25	27	30	33	30	33
Hungary	481	476 - 485	24	28	29	34	29	34
Murcia (Spain)	479	468 - 490						
Russia	478	472 - 483			30	37	30	36
Comunidad Valenciana (Spain)	478	469 - 486						
Luxembourg	477	474 - 479	27	29	32	36	32	36
Iceland	475	472 - 479	28	30	33	37	33	37
Toscana (Italy)	475	467 - 483						
Extremadura (Spain)	473	462 - 485						
Croatia	472	467 - 478			33	40	33	39
Belarus	471	466 - 476			34	40	34	39
Andalusia (Spain)	471	462 - 480						
Canary Islands (Spain)	470	461 - 478						
Ukraine	469	463 - 475			35	42		
Turkey	468	464 - 472	30	32	36	41	36	40
Italy	468	463 - 473	30	33	36	42	36	41
Slovak Republic	464	460 - 469	30	33	39	42	38	41
Israel	462	455 - 469	30	33	38	43	38	42
Malta	457	453 - 460			42	44	41	43
CABA (Argentina)	455	444 - 465						
Sardegna (Italy)	452	444 - 460						
Greece	452	445 - 458	34	35	43	45	42	44
Bogotá (Colombia)	451	441 - 460						
Chile	444	439 - 448	35	35	44	47	43	46
Serbia	440	434 - 446			45	49	44	48
DI Yogyakarta (Indonesia)	439	429 - 449						
Cyprus	439	436 - 442			45	48	44	47
Melilla (Spain)	439	424 - 454						
Malaysia	438	432 - 443			45	50	44	48
United Arab Emirates	434	430 - 438			47	52	47	50
Brunei Darussalam	431	429 - 433			49	53	48	50
Almaty (Kazakhstan)	431	414 - 447						
Jordan	429	424 - 435			49	56		
Moldova	428	424 - 433			49	55		
Astana (Kazakhstan)	428	413 - 443						
DKI Jakarta (Indonesia)	428	415 - 441						
Karagandy region (Kazakhstan)	428	414 - 442						

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028330

• • •

Table I.4.6 [3/3] **Science performance at national and subnational levels**

	Science scale							
			Range of ranks					
			OECD countries		All countries/economies		Countries/economies assessing students on computers	
	Mean score	95% confidence interval	Upper rank	Lower rank	Upper rank	Lower rank	Upper rank	Lower rank
Córdoba (Argentina)	427	418 - 437						
Kostanay region (Kazakhstan)	426	415 - 438						
Thailand	426	420 - 432			50	58	49	54
Uruguay	426	421 - 431			51	57	49	53
Romania	426	417 - 435			49	60		
Bulgaria	424	417 - 431			50	59	49	55
South (Brazil)	419	408 - 431						
Mexico	419	414 - 424	36	37	55	62	51	57
North-Kazakhstan region (Kazakhstan)	419	409 - 429						
Qatar	419	417 - 421			56	60	52	56
Albania	417	413 - 421			57	63	53	58
Costa Rica	416	409 - 422			56	63	52	58
Middle-West (Brazil)	415	399 - 431						
Ceuta (Spain)	415	402 - 428						
Montenegro	415	413 - 418			58	63	54	58
Southeast (Brazil)	414	408 - 419						
PBA (Argentina)	413	403 - 424						
East-Kazakhstan region (Kazakhstan)	413	402 - 424						
Colombia	413	407 - 419	36	37	58	64	54	59
Pavlodar region (Kazakhstan)	413	401 - 425						
North Macedonia	413	410 - 416			60	63		
Peru	404	399 - 409			63	67	58	61
Argentina	404	398 - 410			63	68		
Brazil	404	400 - 408			64	67	59	61
Akmola region (Kazakhstan)	401	391 - 411						
Bosnia and Herzegovina	398	393 - 404			65	70	60	64
Baku (Azerbaijan)	398	393 - 402			66	70	60	64
Zhambyl region (Kazakhstan)	397	389 - 406						
Kazakhstan	397	394 - 400			67	70	61	64
Indonesia	396	391 - 401			67	70	61	64
West-Kazakhstan region (Kazakhstan)	391	381 - 401						
Tucumán (Argentina)	391	381 - 401						
Aktobe region (Kazakhstan)	389	379 - 399						
Saudi Arabia	386	381 - 392			71	73		
North (Brazil)	384	373 - 396						
Lebanon	384	377 - 391			71	74		
Georgia	383	378 - 387			71	74	65	66
Northeast (Brazil)	383	375 - 390						
Almaty region (Kazakhstan)	380	371 - 390						
Morocco	377	371 - 382			73	74	65	66
Kyzyl-Orda region (Kazakhstan)	374	365 - 384						
South-Kazakhstan region (Kazakhstan)	373	366 - 380						
Kosovo	365	363 - 367			75	76	67	68
Panama	365	359 - 370			75	77	67	69
Mangistau region (Kazakhstan)	365	355 - 374						
Atyrau region (Kazakhstan)	361	350 - 371						
Philippines	357	351 - 363			76	77	68	69
Dominican Republic	336	331 - 341			78	78	70	70

1. Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

Notes: OECD countries are shown in bold black. Partner countries, economies and subnational entities that are not included in national results are shown in bold blue. Regions are shown in black italics (OECD countries) or blue italics (partner countries).

Range-of-rank estimates are computed based on mean and standard-error-of-the-mean estimates for each country/economy, and take into account multiple comparisons amongst countries and economies at similar levels of performance. For an explanation of the method, see Annex A3.

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

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028330

Notes

1. Because the membership of the OECD has changed over time, the three categories (around, above and below the OECD mean) are not comparable to the corresponding categories used in earlier PISA reports.
2. See Annex A5 for a discussion of how the scales are linked, and of the comparability of results between paper- and computer-based assessments.
3. While score points in reading, mathematics and science are not comparable, differences in scores can be compared through a standardised effect-size metric, such as Cohen's *d*.
4. In reading, 220 points is approximately equal to the distance between the mid-point of Proficiency Level 5 – a level at which students can 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 – and the mid-point of Proficiency Level 2 – a level at which students are capable of identifying the main idea in a text of moderate length, of finding information based on explicit though sometimes complex criteria, and of reflecting on the purpose and form of texts only when explicitly directed to do so, but have difficulty with reading tasks that do not contain explicit cues or that do contain distractors and competing information (see Chapter 5 for more detailed descriptions of what students can do at different levels of the reading scale).
5. In reading, students in Singapore who reported that they do not speak English at home scored 54 points (S.E.: 3.3 points) below students who reported that they speak English at home; in mathematics, the difference was only 32 points (S.E.: 2.9 points).
6. In this report, the range of ranks is defined as the 97.5% confidence interval for the rank statistic. This means that there is at least a 97.5% probability that the interval defined by the upper and lower ranks, and computed based on PISA samples, contains the true rank of the country/economy (see Annex A3).
7. The lowest rank of country/economy A is not merely given by the number of countries/economies whose mean scores are above those of country/economy A in Table I.4.1, Table I.4.2, and Table I.4.3, and whose names are not listed amongst the non-significant differences compared to country/economy A in those tables. For more details about the methodology behind the computation of a confidence interval for the rank, see Annex A3.
8. In addition to adjudicated subnational entities, whose data were carefully reviewed against technical and scientific standards, the table also includes any subnational entity that constituted one or more explicit sampling strata and that achieved, through deliberate over-sampling or sometimes, due to its large size within the country, a sample of at least 25 participating schools and 875 assessed students. It also includes some subnational entities that conducted a census, and where the country requested that results be reported at the subnational level. For non-adjudicated entities, response rates were not assessed separately from those of the country as a whole, and results must be interpreted with caution.
9. If the distribution of performance amongst the eligible 15-year-olds (first-order) stochastically dominates that of the non-eligible 15-year-olds, then the mean and all percentiles of the PISA target population represent an upper bound on the percentiles of the population encompassing all 15-year-olds.
10. See https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (accessed on 23 August 2019).
11. The GDP values represent per capita GDP in 2018 at current prices, expressed in USD. The conversion from local currencies to equivalent USD accounts for differences in purchasing power across countries and economies.
12. Spending per student is approximated by multiplying the expenditure per student on educational institutions in 2018 (from public and private sources), at each level of education, by the theoretical duration of education at the respective level, up to the age of 15. Cumulative expenditure for a given country is approximated as follows: let n_0, n_1 and n_2 be the typical number of years spent by a student from the age of 6 up to the age of 15 in primary, lower secondary and upper secondary education. Let E_0, E_1 and E_2 be the annual expenditure per student in USD converted using purchasing power parity in primary, lower secondary and upper secondary education, respectively. The cumulative expenditure is then calculated by multiplying current annual expenditure for each level of education by the typical duration of study in that level, using the following formula: $CE=n_0E_0+n_1E_1+n_2 E_2$.
13. The countries and economies included in each analysis may vary due to data availability. The percentage of variation in mean reading performance accounted for by each variable cannot therefore be directly compared.
14. The indicator of total learning time computed based on 2015 data is used as a proxy for the time investment of PISA 2018 students, because PISA 2018 did not collect data on out-of-school learning time.
15. Different countries participated in the Survey of Adult Skills (PIAAC) in different years. In all countries, results for 35-54 year-olds are approximated by the results of adults born between 1964 and 1983. No adjustment is made to account for changes in the skills of these adults, or for changes in the composition of these cohorts, between the year in which the survey was conducted and 2018. PISA results for the Flemish Community of Belgium are related to PIAAC results for Flanders (Belgium). PIAAC results for Ecuador are related to the country's results in the PISA for Development assessment (2017). For the United States, PIAAC data refer to 2017.
16. PISA and PIRLS assess different constructs and different samples. For example, PIRLS uses a grade-based definition of the target population, while PISA uses an age-based definition. Dropout between the end of primary school and the age of 15 may reduce the comparability of samples across assessments. Also note that the cohort that was assessed in PIRLS 2011 differs by 1 or 2 years, in most cases, from the cohort assessed in PISA 2018. In addition, cohort composition could have changed in some countries and economies due to migration. It is beyond the scope of this chapter to analyse these differences in detail.

17. As noted in Worden (2012[9]), bilingualism and multilingualism can have multiple benefits for students and should be encouraged. Bilingualism, in particular, is associated with enhanced executive control (Bialystok, 2011[11]). Despite the many advantages of bilingualism, it has been shown that bilingual children, on average, know significantly fewer words in each language than comparable monolingual children (Bialystok et al., 2009[10]). Several high-performing countries in PISA have large shares of bilingual students, including Singapore, one of the highest-performing countries in all subjects, and Switzerland, which scores around the OECD average in reading, but above the OECD average in mathematics.

18. International PISA data cannot describe all aspects of ethnic diversity. For example, in Australia, New Zealand or in the Americas, PISA measures of linguistic diversity and immigrant status do not necessarily cover indigenous populations, which use the language of instruction in everyday life.

References

Bialystok, E. (2011), "Reshaping the mind: The benefits of bilingualism.", *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, Vol. 65/4, pp. 229-235, http://dx.doi.org/10.1037/a0025406. [11]

Bialystok, E. et al. (2009), "Bilingual Minds", *Psychological Science in the Public Interest*, Vol. 10/3, pp. 89-129, http://dx.doi.org/10.1177/1529100610387084. [10]

Mullis, I. et al. (2012), *PIRLS 2011 International Results in Reading*, TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA), https://timssandpirls.bc.edu/pirls2011/downloads/P11_IR_FullBook.pdf (accessed on 3 July 2019). [5]

OECD (2019), *Education at a Glance 2019: OECD Indicators,* OECD Publishing, Paris, https://doi.org/10.1787/f8d7880d-en. [6]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [1]

OECD (2019), *The Road to Integration: Education and Migration*, OECD Reviews of Migrant Education, OECD Publishing, Paris, https://dx.doi.org/10.1787/d8ceec5d-en. [8]

OECD (2016), *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264267510-en. [4]

OECD (2010), *Educating Teachers for Diversity: Meeting the Challenge*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264079731-en. [7]

Ward, M. (2018), "PISA for Development: Results in Focus"*, PISA in Focus*, No. 91, OECD Publishing, Paris, https://dx.doi.org/10.1787/c094b186-en. [2]

Worden, J. (2012), "Bilingual education policy and language learning in Estonia and Singapore", in *Languages in a Global World: Learning for Better Cultural Understanding*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264123557-11-en. [9]

World Bank (2017), *World Development Report 2018: Learning to Realize Education's Promise*, The World Bank, http://dx.doi.org/ 10.1596/978-1-4648-1096-1. [3]

5

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.

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%	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. 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.
5	626	8.7%	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. 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. 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.
4	553	27.6%	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 *ad hoc* categories. They can compare perspectives and draw inferences based on multiple sources. 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. 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. 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.

...

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%	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. 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. 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. 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.
2	407	77.4%	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. 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. 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.
1a	335	92.3%	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. 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. 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. 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.
1b	262	98.6%	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. 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. 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.
1c	189	99.9%	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. Tasks at this level involve simple vocabulary and syntactic structures.

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)	*323*
		Most reading-fluency tasks calling for a "no" response (meaningless sentences, such as "Airplanes are made of dogs")	
1c	189	*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*	

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.

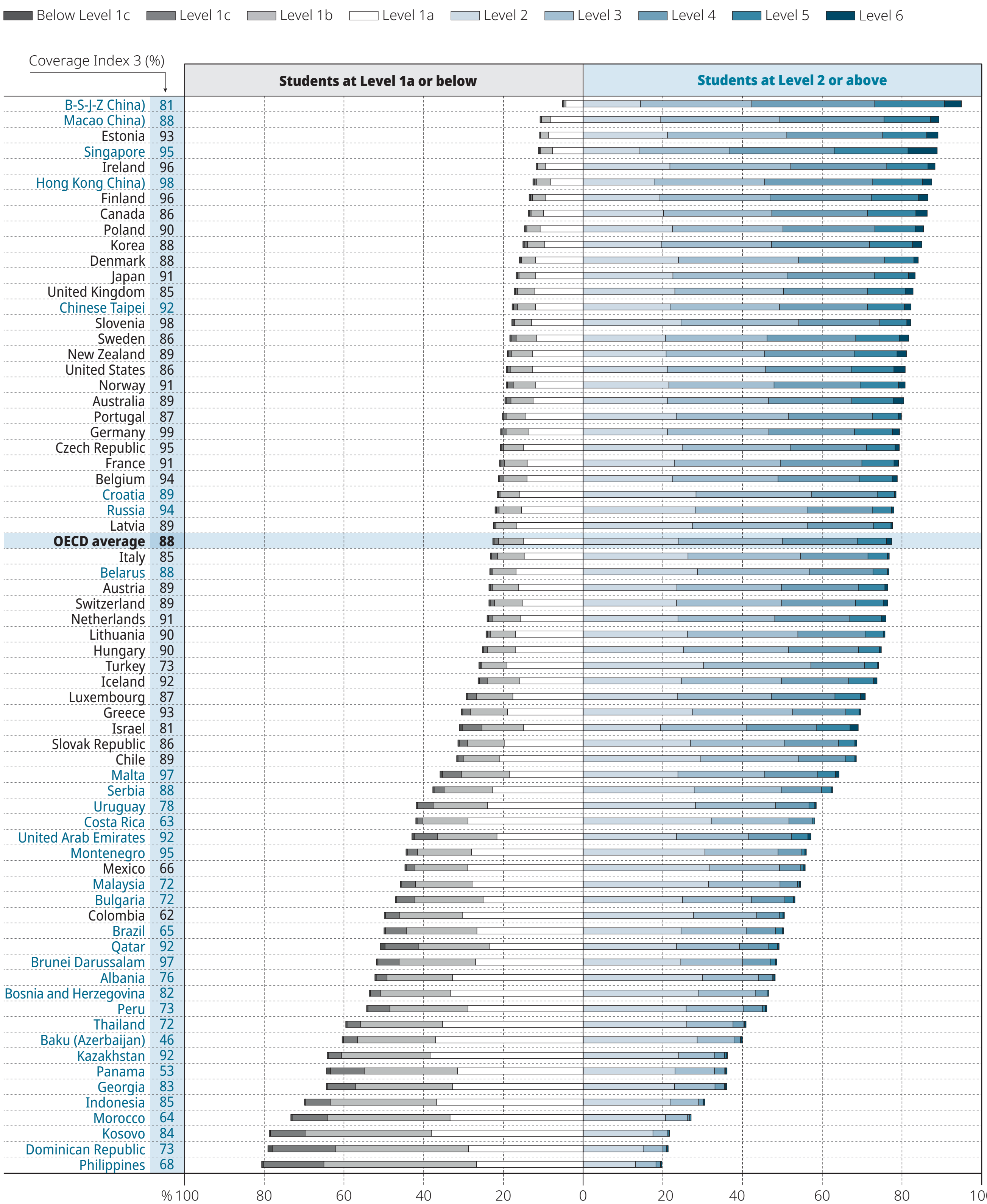
PERCENTAGE OF STUDENTS AT THE DIFFERENT LEVELS OF READING PROFICIENCY

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).

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



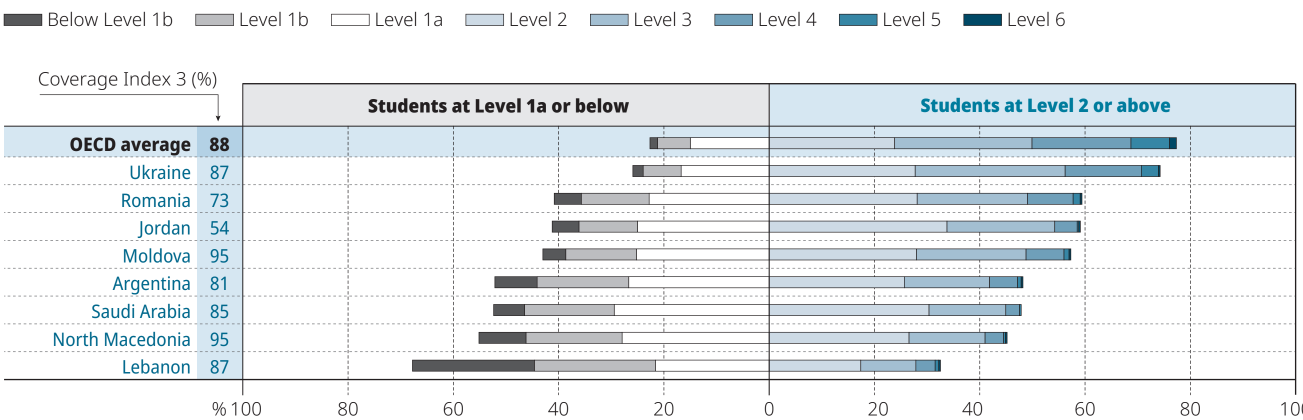
Note: Coverage Index 3 is shown next to the country/economy name.

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 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 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 *COW'S 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]

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

Notes

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, https://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]

6

What can students do in mathematics?

This chapter describes the range of mathematical competences assessed in PISA 2018 and reports the proportion of students who performed at each level of proficiency.

The PISA assessment of mathematics focuses on measuring students' capacity to formulate, use and interpret mathematics in a variety of contexts. These include not only familiar settings related to personal experience, such as when preparing food, shopping or watching sports, but also occupational, societal and scientific contexts, such as costing a project, interpreting national statistics or modelling natural phenomena. To succeed on the PISA test, students must be able to reason mathematically and use mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. Competence in mathematics, as defined in PISA, assists individuals in recognising the role that mathematics plays in the world and in making the well-founded judgements and decisions needed to be constructive, engaged and reflective citizens (OECD, 2019[1]).[1]

Performance in mathematics described in this way encompasses more than the ability to reproduce the mathematics concepts and procedures acquired in school. PISA seeks to measure how well students can extrapolate from what they know and apply their knowledge of mathematics in a range of situations, including new and unfamiliar ones. To this end, most PISA mathematics units refer to real-life contexts in which mathematics abilities are required to solve a problem. The focus on real-life contexts is also reflected in the possibility of using "tools", such as a calculator, a ruler or a spreadsheet, to solve problems, just as one would do in a real-life situation.

What the data tell us

- On average across OECD countries, 76% of students attained Level 2 or higher in mathematics. At a minimum, these students can interpret and recognise, without direct instructions, how a (simple) situation can be represented mathematically (e.g. comparing the total distance across two alternative routes, or converting prices into a different currency). However, in 24 countries and economies, more than 50% of students scored below this level of proficiency.
- Six Asian countries and economies had the largest shares of students who scored at Level 5 or higher in mathematics: Beijing, Shanghai, Jiangsu and Zhejiang (China) (44%), Singapore (37%), Hong Kong (China) (29%), Macao (China) (28%), Chinese Taipei (23%) and Korea (21%). These students can model complex situations mathematically, and can select, compare and evaluate appropriate problem-solving strategies for dealing with them.
- About one in six 15-year-old students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (16%), and about one in seven students in Singapore (14%), scored at Level 6 in mathematics, the highest level of proficiency that PISA describes. These students are capable of advanced mathematical thinking and reasoning. On average across OECD countries, only 2.4% of students scored at this level.
- Compared to countries of similar average performance in PISA, Germany and Korea have a larger share of students who performed at the highest levels of mathematics proficiency, but also of students who performed at the lowest levels. This reflects the wide variation in mathematics performance within these countries.

Mathematics was tested using computers (as were reading and science) in 70 of the 79 participating countries and economies; the remaining 9 countries delivered the test in a pencil-and-paper format. All countries/economies, regardless of the assessment mode, used the same mathematics questions.[2] Results of the PISA mathematics test can be compared across all 79 participating countries and economies. Annex A5 discusses the differences between paper- and computer-based assessments. It explains how, in order to report results on the same scale and enable fair comparisons, 32 (out of 82) items were deemed to vary in difficulty between the computer-based and paper-and-pencil tests (6 items were deemed easier on computer; 26 items were deemed more difficult).

THE RANGE OF PROFICIENCIES COVERED BY THE PISA MATHEMATICS TEST

As discussed in Chapter 2, student performance in PISA is reported on a scale. To help interpret what students' scores mean in substantive terms, the scale is divided into levels of proficiency that indicate the kinds of tasks that students whose scores are above a lower score limit are capable of completing successfully. The six proficiency levels used in the PISA 2018 mathematics assessment were the same as those established for the PISA 2003 and 2012 assessments, when mathematics was the major area of assessment. The process used to produce proficiency levels in mathematics is described in Chapter 2. Table I.6.1 illustrates the range of mathematical skills that are covered by the PISA test and describes the skills, knowledge and understanding that are required at each level of the mathematics scale.

Since it is necessary to preserve the confidentiality of the test material in order to continue to monitor trends in mathematics beyond 2018, the questions used in the PISA 2018 assessment of mathematics cannot be presented in this report. However, it is still possible to illustrate the proficiency levels with questions from previous assessments. Sample items that illustrate the different levels of mathematics proficiency can be found in *PISA 2012 Results: What Students Know and Can Do* (OECD, 2014, pp. 125-142[2]).

Table I.6.1 **Summary description of the six levels of mathematics 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	669	2.4%	At Level 6, students can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations together and flexibly translate amongst them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop new approaches and strategies for attacking novel situations. Students at this level can reflect on their actions, and can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments and the appropriateness of these to the original situation.
5	607	10.9%	At Level 5, students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. Students at this level have begun to develop the ability to reflect on their work and to communicate conclusions and interpretations in written form.
4	545	29.5%	At Level 4, students can work effectively with explicit models for complex, concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic representations, linking them directly to aspects of real-world situations. Students at this level can utilise their limited range of skills and can reason with some insight in straightforward contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments and actions.
3	482	53.8%	At Level 3, students can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning.
2	420	76.0%	At Level 2, students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures or conventions to solve problems involving whole numbers. They are capable of making literal interpretations of results.
1	358	90.9%	At Level 1, students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are almost always obvious and follow immediately from the given stimuli.

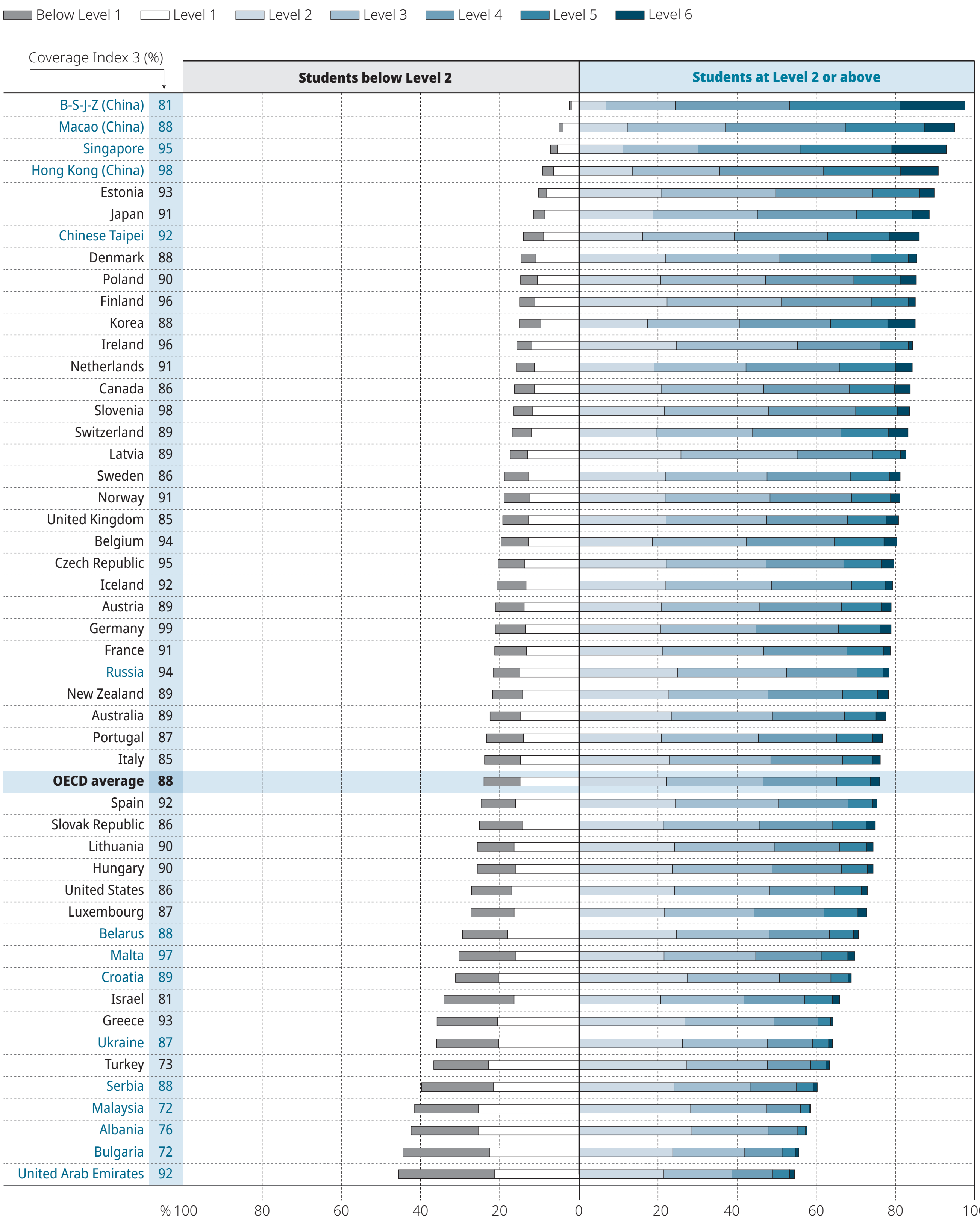
PERCENTAGE OF STUDENTS AT THE DIFFERENT LEVELS OF MATHEMATICS PROFICIENCY

Figure I.6.1 shows the distribution of students across the six proficiency levels in each participating country and economy. Table I.B1.2 (in Annex B1) shows these same percentages of students at each proficiency level on the mathematics scale, with standard errors.

Proficiency at Level 2 or above

This report considers students who scored below Level 2 as "low-achieving students". Indeed, at Level 2, students begin to demonstrate the ability and initiative to use mathematics in simple real-life situations. The global indicators for the United Nations Sustainable Development Goals (see Chapter 11) identify Level 2 proficiency as the "minimum level of proficiency" that all children should acquire by the end of secondary education. While students who score below this minimum level can be considered particularly at risk, Level 2 proficiency is by no means a "sufficient" level of mathematics proficiency for making well-founded judgements and decisions across a range of personal or professional situations in which mathematical literacy is required.

Figure I.6.1 [1/2] **Students' proficiency in mathematics**



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

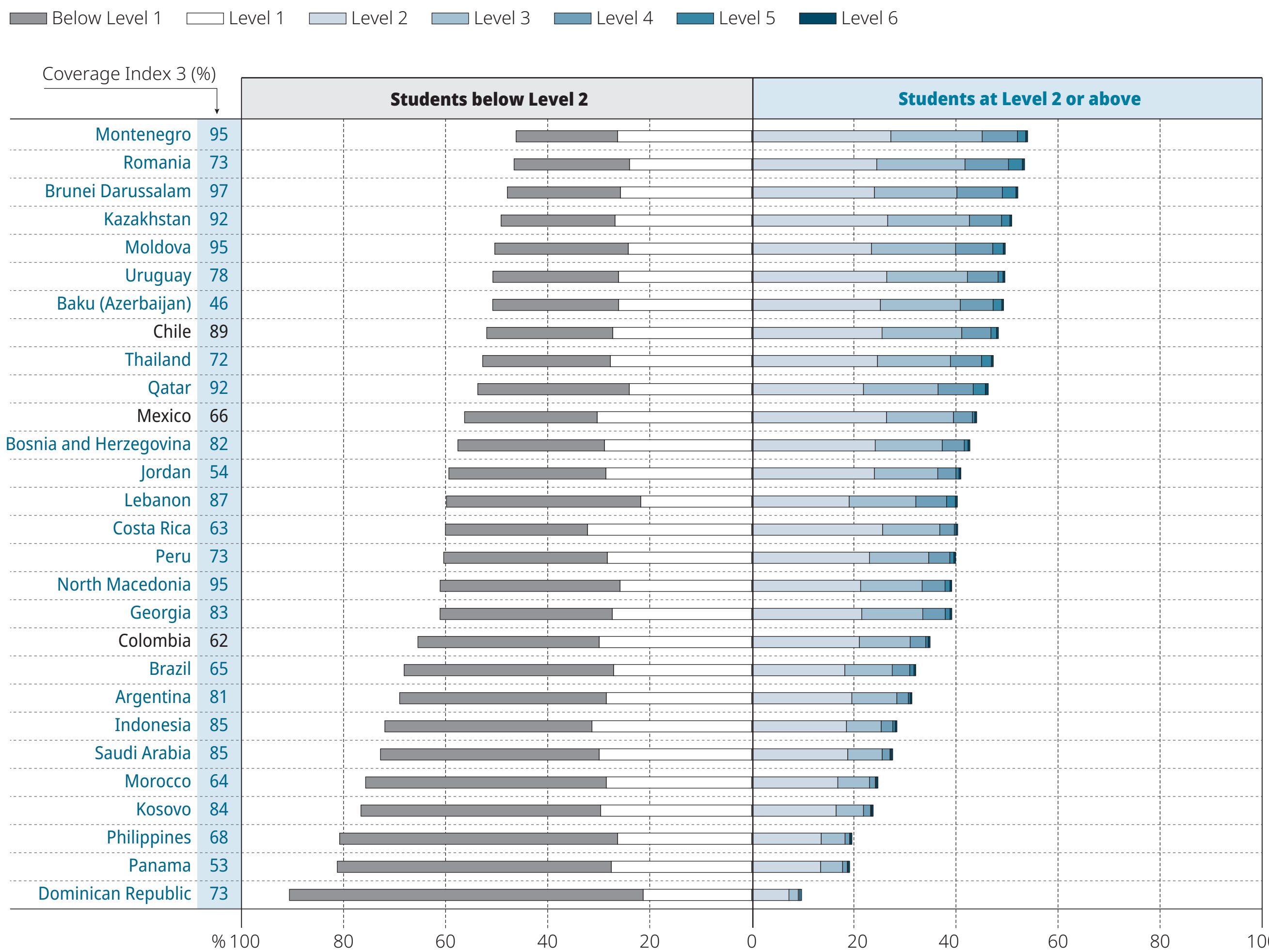
Countries and economies are ranked in descending order of the percentage of students who performed at or above Level 2.

Source: OECD, PISA 2018 Database, Tables I.B1.2 and I.A2.1.

StatLink https://doi.org/10.1787/888934028634

...

Figure I.6.1 [2/2] **Students' proficiency in mathematics**



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.B1.2 and I.A2.1.

StatLink https://doi.org/10.1787/888934028634

Skills requirements are likely to evolve over time and to depend on the context and the tools, such as technologies, that one can make use of in that context. As more advanced technologies that can substitute for certain human skills become available, the skills required for participation in labour markets, for example, are likely to increase (Goldin and Katz, 2008[3]; Elliott, 2017[4]; Frey and Osborne, 2017[5]).

Proficiency at Level 2

At Level 2, students can use basic algorithms, formulae, procedures or conventions to solve problems involving whole numbers – e.g. to compute the approximate price of an object in a different currency or to compare the total distance across two alternative routes. They can interpret and recognise situations in contexts that require no more than direct inference, extract relevant information from a single source and make use of a single representational mode (such as graphs, tables, equations, etc). Students at this level are capable of making literal interpretations of results.

More than 90% of students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]), Hong Kong (China), Macao (China) and Singapore, and close to 90% in Estonia achieved this benchmark. On average across OECD countries, 76% of students attained Level 2 or higher (that is, were proficient at Level 2, 3, 4, 5 or 6; Figure I.6.1 and Table I.B1.2). Meanwhile, fewer than one in ten students in the Dominican Republic (9.4%), and only 19% of students in Panama and the Philippines attained this baseline level of mathematics proficiency. In 21 other countries, more than 20% but less than 50% of 15-year-old students attained this level of proficiency.

Proficiency at Level 3

At Level 3, students can execute clearly described procedures, including those that require sequential decisions. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their interpretations are sufficiently sound to be the basis for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. Their solutions show that they have engaged in basic interpretation and reasoning.

Across OECD countries, 54% of students were proficient at Level 3 or higher. More than 90% of students were proficient at Level 3 or higher in B-S-J-Z (China); and at least two out of three students in Estonia, Hong Kong (China), Japan, Korea, Macao (China), Singapore and Chinese Taipei attained this level. This shows that in some education systems, virtually all children develop a solid foundation in mathematics. In contrast, in 24 countries and economies with comparable data, three out of four students did not attain this level (Figure I.6.1 and Table I.B1.2).

Proficiency at Level 4

At Level 4, students can work effectively with explicit models on complex, concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic representations (such as equations and formulae), linking them directly to aspects of real-world situations. Students at this level can reason with some insight. They can construct and communicate explanations and arguments based on their interpretations, reasoning and actions.

Across OECD countries, 29% of students performed at proficiency Level 4, 5 or 6. More than one in two students in B-S-J-Z (China) (73%), Singapore (63%), Macao (China) (58%) and Hong Kong (China) (55%) performed at one of these levels. Between 40% and 50% of students in Chinese Taipei (47%), Korea (44%), Japan (43%), the Netherlands (42%) and Estonia (40%) scored at or above Level 4. By contrast, in 27 participating countries and economies with comparable data – including all of the Latin American countries that participated in PISA 2018 – less than 10% of students attained this level (Figure I.6.1 and Table I.B1.2).

Proficiency at Level 5

At Level 5, students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insights pertaining to these situations. They have begun to develop the ability to reflect on their work, and to communicate conclusions and interpretations in written form.

Across OECD countries, 11% of students were top performers in 2018, meaning that they were proficient at Level 5 or 6. Amongst all countries and economies that participated in PISA 2018, B-S-J-Z (China) had the largest proportion of top performers (44%), followed by five other Asian countries and economies: Singapore (37%), Hong Kong (China) (29%), Macao (China) (28%), Chinese Taipei (23%) and Korea (21%). The strong performance of these Asian countries dispels the notion that children in these countries merely memorise subject-matter content: Level 5 proficiency requires students to master a high level of conceptual understanding and mathematical reasoning. In all remaining countries and economies, less than 20% of students attained this level of proficiency in mathematics.

Countries with similar mean performance may have significantly different shares of students who are able to perform at the highest levels in PISA (Figure I.6.1 and Table I.B1.2). A smaller share of top-performing students, compared to countries of similar average performance, means that student performance varies more narrowly around the mean. For example:

- Estonia (mean score: 523 points) had significantly fewer top-performing students (15.5%) compared to Japan and Korea (mean scores: 527 and 526 points, respectively; top performers: 18.3% and 21.4%, respectively)
- Germany and Ireland performed similarly, on average (mean score of 500 points in both countries), but 13.3% of students in Germany were top performers, compared to just 8.2% in Ireland
- Croatia and Israel also performed similarly, on average (mean scores of 464 points and 463 points, respectively); yet 5.1% of students in Croatia were top performers compared to 8.8% of students in Israel.

Proficiency at Level 6

Students at Level 6 on the PISA mathematics scale can successfully complete the most difficult PISA items. These students can conceptualise, generalise and use information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations together and move flexibly amongst them. Students at this level are capable of advanced mathematical thinking and reasoning.

These students can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop approaches and strategies for addressing novel situations. Students at this level can reflect on their actions, can formulate and precisely communicate those actions and reflections and can explain why they were applied to the original situation.

On average across OECD countries, only 2.4% of students attained Level 6. About one in six students scored at this level in B-S-J-Z (China) (16%), and about one in seven students in Singapore (14%). In Hong Kong (China), Korea, Macao (China) and Chinese Taipei, between 5% and 10% of students attained proficiency Level 6. In 36 participating countries and economies, between 1% and 5% of students performed at this level; in 21 countries/economies, between 0.1% and 1% of students scored at Level 6; and in 15 other countries/economies, fewer than 1 in 1 000 students (0.1%) performed at Level 6 (Figure I.6.1 and Table I.B1.2).

Proficiency below Level 2

The PISA 2018 mathematics assessment identified one proficiency level below Level 2. Students who scored at or below this level are considered low achievers in mathematics.

Proficiency at Level 1

At Level 1, students can answer mathematics questions involving familiar contexts where all of the relevant information is present and the questions are clearly defined. They are able to identify information and carry out routine procedures according to direct instructions. They can only perform actions that are obvious and that follow immediately from the given stimuli.

On average across OECD countries in 2018, 15% of students were proficient only at Level 1. In B-S-J-Z (China) (2.4%), Macao (China) (5.0%), Singapore (7.1%) and Hong Kong (China) (9.2%), less than 10% of students performed at or below Level 1 (Figure I.6.1 and Table I.B1.2). But in 21 countries and economies, Level 1 was the median level of proficiency, meaning that the score which divides 15-year-old students into two equal halves – 50% scoring above it, and 50% scoring below it – fell within the range of Level 1.

Proficiency below Level 1

Some 9.1% of students on average across OECD countries scored below Level 1 in mathematics. By contrast, in the Dominican Republic (69%), the Philippines (54%) and Panama (54%), more than one in two students scored below Level 1, the lowest described level of proficiency in PISA. In 26 participating countries and economies, between 20% and 50% of students did not reach Level 1 on the mathematics scale.

The PISA mathematics test included too few tasks of the appropriate difficulty that would help describe an additional level of proficiency below Level 1. However, based on the few PISA 2012 mathematics items whose difficulty lies below Level 1 (four of which were also included in the PISA 2018 mathematics assessment), students who score below Level 1, but not too far from it, can be expected to perform some direct and straightforward mathematical tasks. These include reading a single value from a well-labelled chart or table, where the labels on the chart match the words in the stimulus and question, so that the selection criteria are clear and the relationship between the chart and the aspects of the context depicted are evident. They may also be able to perform simple arithmetic calculations with whole numbers by following clear and well-defined instructions.

Given the large number of students who scored at these levels in many PISA-participating countries, the group of international experts working on the PISA 2021 test is trying to broaden the range of fundamental mathematical capabilities that PISA assesses, based partly on approaches piloted in the PISA for Development project (OECD, 2018[6]).

All PISA-participating countries and economies have students who score at or below Level 1, but the largest proportions of students who score at these levels are found in the lowest-performing countries. In some cases, countries with similar mean performance may have significantly different shares of students who score below Level 2 in mathematics. For example, in Estonia, whose mean performance in 2018 (523 score points) was not significantly different from that of Korea or the Netherlands, only 10.2% of students scored at or below Level 1, while 15.0% did so in Korea and 15.8% in the Netherlands. This shows that in Korea and the Netherlands, student performance in mathematics varied more widely than in Estonia, despite similar average performance. And while mean performance in Germany and Ireland was similar (500 score points), the percentage of low achievers in Germany (21.1%) was about 5 percentage points higher than that in Ireland (15.7%).

ACCOUNTING FOR OUT-OF-SCHOOL 15-YEAR-OLDS

When evaluating countries' success in equipping young adults with solid reading, mathematics or science skills, it is also important to consider whether these comparisons may change if 15-year-olds who are not part of the PISA target population were also included. For this reason, Figure I.6.1 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).[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 were not represented by the PISA sample would have scored had they taken the assessment. However, for countries where many 15-year-olds were not enrolled or were 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 (see Box I.5.1 in Chapter 5). Accounting for changing rates of out-of-school 15-year-olds is particularly important when comparing countries' performance over time (see Chapter 9), or when assessing countries' performance against global development goals for the education of all children (see Chapter 10).

Notes

1. Mathematics was the major domain assessed in 2003 and 2012. In the 2018 PISA assessment, an update on overall performance in mathematics is presented. For more in-depth analyses of the mathematics assessment and student performance in mathematics, see the reports based on data from PISA 2003 and PISA 2012 (OECD, 2004[7]; OECD, 2010[8]; OECD, 2014[2]; OECD, 2016[9]; Echazarra et al., 2016[10]).

2. Only one item included in the PISA 2018 paper-based assessment was not included in the PISA 2018 computer-based assessment. As a result, the total number of items in the paper-based assessment is 83, while the total number of items in the computer-based assessment is 82 (see Annex A5 for details).

3. While the number of 15-year-olds who are covered by the PISA sample is estimated from school-enrolment estimates (for all schools) and from student lists (for sampled schools), the total population of 15-year-olds is based on demographic projections provided by PISA national centres. The difference between the two numbers may also reflect errors in the projections. Fifteen-year-olds who are not represented by PISA samples also include students and schools that were excluded from sampling (see Chapter 3 and Annex A2) and a small fraction of students who were in the process of transferring between schools at the time of the PISA test.

References

Echazarra, A. et al. (2016), "How teachers teach and students learn: Successful strategies for school", *OECD Education Working Papers*, No. 130, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jm29kpt0xxx-en. [10]

Elliott, S. (2017), *Computers and the Future of Skill Demand*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264284395-en. [4]

Frey, C. and **M. Osborne** (2017), "The future of employment: How susceptible are jobs to computerisation?", *Technological Forecasting and Social Change*, Vol. 114, pp. 254-280, http://dx.doi.org/10.1016/j.techfore.2016.08.019. [5]

Goldin, C. and **L. Katz** (2008), *The Race between Education and Technology*, Belknap Press of Harvard University Press. [3]

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [1]

OECD (2018), "PISA for Development Mathematics Framework", in *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264305274-5-en. [6]

OECD (2016), *Equations and Inequalities: Making Mathematics Accessible to All*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264258495-en. [9]

OECD (2014), *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, https://dx.doi.org/10.1787/9789264208780-en. [2]

OECD (2010), *Mathematics Teaching and Learning Strategies in PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264039520-en. [8]

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264006416-en. [7]

7

What can students do in science?

This chapter describes the range of science competences assessed in PISA 2018 and reports the proportion of students who performed at each level of proficiency.

The PISA assessment of science focuses on measuring students' ability to engage with science-related issues and with the ideas of science, as reflective citizens. Engaging in reasoned discourse about science and science-based technology requires a sound knowledge of facts and theories to explain phenomena scientifically. It also requires knowledge of the standard methodological procedures used in science, and knowledge of the reasons and ideas used by scientists to justify their claims, in order to evaluate (or design) scientific enquiry and to interpret evidence scientifically.

In contemporary societies, an understanding of science and of science-based technology is necessary not only for those whose careers depend on it directly, but also for any citizen who wishes to make informed decisions related to the many controversial issues under debate today – from personal issues, such as maintaining a healthy diet, to local issues, such as how to manage waste in big cities, to global and far-reaching issues, such as the costs and benefits of genetically modified crops or how to prevent and mitigate the negative consequences of global warming on physical, ecological and social systems.

What the data tell us

- On average across OECD countries, 78% of students attained Level 2 or higher in science. At a minimum, these students can recognise the correct explanation for familiar scientific phenomena and can use such knowledge to identify, in simple cases, whether a conclusion is valid based on the data provided. More than 90% of students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (97.9%), Macao (China) (94.0%), Estonia (91.2%) and Singapore (91.0%) achieved this benchmark.
- On average across OECD countries, 6.8% of students were top performers in science in 2018, meaning that they were proficient at Level 5 or 6. Almost one in three (32%) students in Beijing, Shanghai, Jiangsu and Zhejiang (China), and more than one in five students in Singapore (21%) performed at this level. In addition to skills associated with lower proficiency levels, these students can creatively and autonomously apply their knowledge of and about science to a wide variety of situations, including unfamiliar ones.

Science was the major domain assessed in both 2006 and 2015. The PISA science test was significantly expanded in 2015 to make use of the capabilities of computers, the new mode of delivery used in most participating education systems. For example, through its interactive interface, PISA 2015 was able, for the first time, to assess students' ability to conduct scientific enquiry by asking test-takers to design (simulated) experiments and interpret the resulting evidence. The main part of this chapter covers the range of science proficiency as assessed in the computer-based test of science.

The nine countries that participated in PISA 2018 using pen-and-paper tests continued to use tasks designed initially for the 2006 assessment. Because some of these tasks were adapted and also used in countries that delivered the science test on computer, results can be reported on the same numeric scale (something that is particularly important for assessing performance trends over time that start from earlier pen-and-paper assessments, including in countries that conducted the PISA 2018 science test on computer). However, strictly speaking, these scores should be interpreted according to different descriptors of proficiency. When describing the performance of students in these nine countries, this chapter therefore also highlights the most relevant distinctions between the range of proficiency assessed through the pen-and-paper test (which does not include the ability to carry out experiments and conduct scientific enquiry) and the wider range assessed through computer delivery of the test.

THE RANGE OF PROFICIENCIES COVERED BY THE PISA SCIENCE TEST

As discussed in Chapter 2, student performance in PISA is reported as a score on a scale. To help interpret what students' scores mean in substantive terms, the scale is divided into levels of proficiency that indicate the kinds of tasks that students at those levels are capable of completing successfully. The seven proficiency levels used in the PISA 2018 science assessment were the same as those established for the PISA 2015 assessment.[1] The process used to produce proficiency levels in science is described in Chapter 2. Table I.7.1 illustrates the range of science competences covered by the PISA test and describes the skills, knowledge and understanding that are required at each level of the science scale.

Since it is necessary to preserve the confidentiality of the test material in order to continue to monitor trends in science beyond 2018, the questions used in the PISA 2018 assessment of science cannot be presented in this report. Instead, it is possible to illustrate the proficiency levels with questions that were released after previous assessments. Sample items that illustrate the different levels of science proficiency can be found in Annex C of *PISA 2015 Results (Volume I)* (OECD, 2016, pp. 462-481[11]) and on line at www.oecd.org/pisa/test/.

Table I.7.1 **Summary description of the seven levels of science 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	708	0.8%	At Level 6, students can draw on a range of interrelated scientific ideas and concepts from the physical, life, and earth and space sciences and use content, procedural and epistemic knowledge in order to offer explanatory hypotheses of novel scientific phenomena, events and processes or to make predictions. In interpreting data and evidence, they are able to discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. Level 6 students can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.
5	633	6.8%	At Level 5, students can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes involving multiple causal links. They are able to apply more sophisticated epistemic knowledge to evaluate alternative experimental designs and justify their choices, and use theoretical knowledge to interpret information or make predictions. Level 5 students can evaluate ways of exploring a given question scientifically and identify limitations in interpretations of data sets, including sources and the effects of uncertainty in scientific data.
4	559	24.9%	At Level 4, students can use more complex or more abstract content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context. They are able to justify an experimental design by drawing on elements of procedural and epistemic knowledge. Level 4 students can interpret data drawn from a moderately complex data set or less familiar context, draw appropriate conclusions that go beyond the data and provide justifications for their choices.
3	484	52.3%	At Level 3, students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 students are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.
2	410	78.0%	At Level 2, students are able to draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data and identify the question being addressed in a simple experimental design. They can use basic or everyday scientific knowledge to identify a valid conclusion from a simple data set. Level 2 students demonstrate basic epistemic knowledge by being able to identify questions that can be investigated scientifically.
1a	335	94.1%	At Level 1a, students are able to use basic or everyday content and procedural knowledge to recognise or identify explanations of simple scientific phenomena. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a students can select the best scientific explanation for given data in familiar personal, local and global contexts.
1b	261	99.3%	At Level 1b, students can use basic or everyday scientific knowledge to recognise aspects of familiar or simple phenomena. They are able to identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.

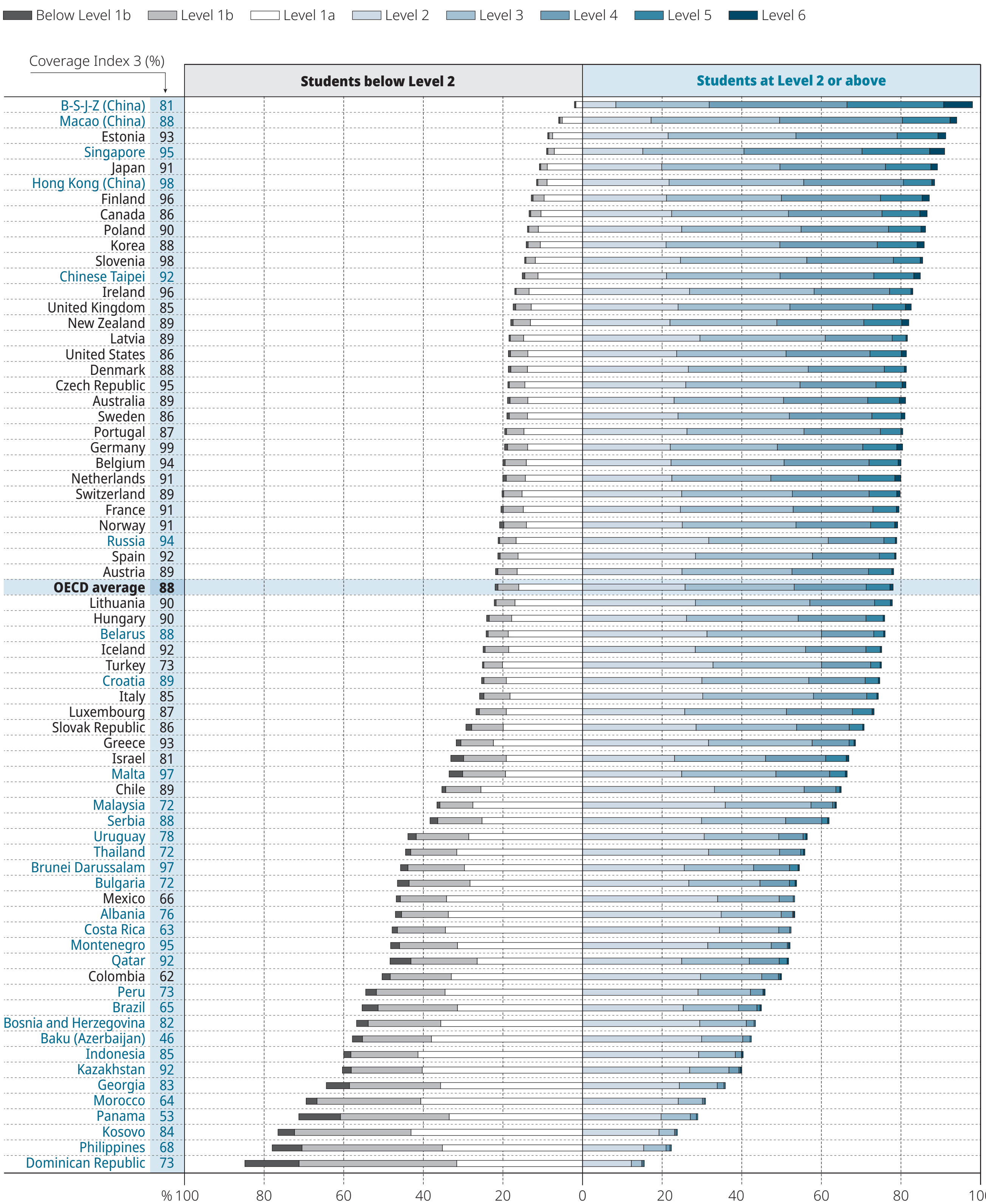
PERCENTAGE OF STUDENTS AT THE DIFFERENT LEVELS OF SCIENCE PROFICIENCY

Figure I.7.1 and, for countries that used paper test booklets, Figure I.7.2, show the distribution of students across the seven proficiency levels in each participating country and economy. Table I.B1.3 (in Annex B1) shows these same percentages of students at each proficiency level on the science scale, with standard errors.

Proficiency at or above Level 2

Level 2 in science is an important benchmark for student performance: it represents the level of achievement, on the PISA scale, at which students begin to demonstrate the science competences that will enable them to engage in reasoned discourse about science and technology (OECD, 2018, p. 72[2]). At Level 2, the attitudes and competences required to engage effectively with science-related issues are only just emerging. Students demonstrate basic or everyday scientific knowledge, and a basic understanding of scientific enquiry, which they can apply mostly in familiar contexts. Students' skills progressively expand to less familiar contexts, and to more complex knowledge and understanding at higher levels of proficiency.

Figure I.7.1 **Students' proficiency in science (computer-based assessment)**



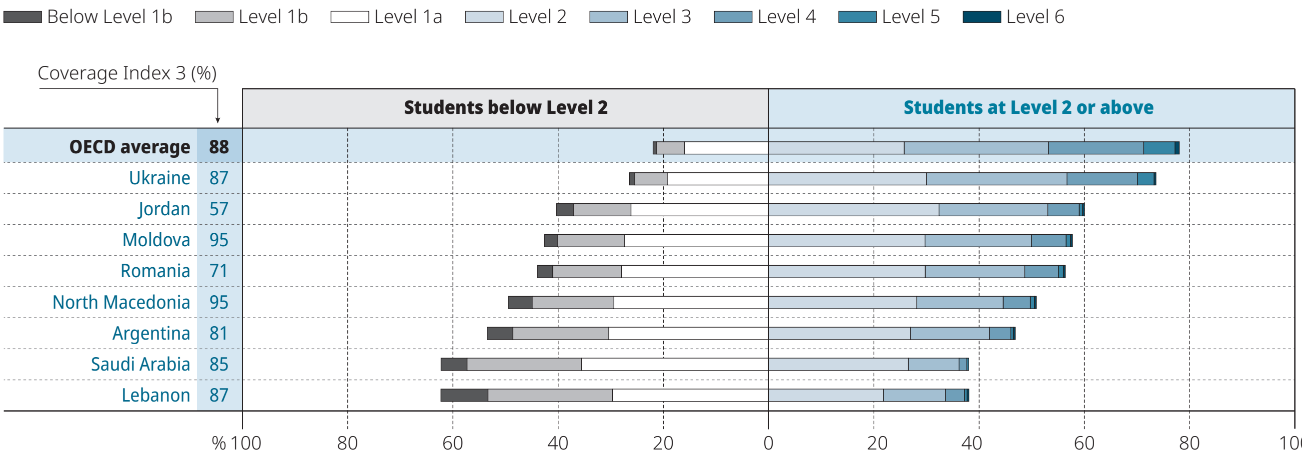
Note: Coverage Index 3 is shown next to the country/economy name.

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.B1.3 and I.A2.1.

StatLink https://doi.org/10.1787/888934028653

Figure I.7.2 **Students' proficiency in science (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.B1.3 and I.A2.1.

StatLink https://doi.org/10.1787/888934028672

Level 2 does not establish a threshold for scientific illiteracy. PISA views science literacy not as an attribute that a student has or does not have, but as a set of skills that can be acquired to a greater or lesser extent. It also does not identify a "sufficient" level of science literacy, particularly not for those whose careers will directly depend on an understanding of science and of science-based technology. However, Level 2 does establish a baseline threshold below which students typically require some support to engage with science-related questions, even in familiar contexts. For this reason, this report describes students performing below Level 2 as "low-achieving students".

Proficiency at Level 2

At Level 2, students can draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data, and identify the question being addressed in a simple experimental design. They can use common scientific knowledge to identify a valid conclusion from a simple data set. Level 2 students demonstrate basic epistemic knowledge by being able to identify questions that could be investigated scientifically.

Level 2 can be considered as the level of science proficiency at which students begin to demonstrate the competences that will enable them to engage effectively and productively with issues related to science and technology. More than 90% of students in Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]") (97.9%), Macao (China) (94.0%), Estonia (91.2%) and Singapore (91.0%) met this benchmark. Across OECD countries, an average of 78% of students attained Level 2. Meanwhile, only about one in six students in the Dominican Republic (15%) and only a minority (less than 50%, but more than 20%) of students in 15 other countries and economies attained this level of proficiency (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Proficiency at Level 3

At Level 3, students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context (the ability to carry out experiments was not assessed in paper-based tests). Level 3 students can distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.

On average across OECD countries, more than half of all students (52%) were proficient at Level 3 or higher (that is, at Level 3, 4, 5 or 6). The average median score across OECD countries, i.e. the score that divides the population in two equal halves (one half scoring above the median, and the other half below), fell within Level 3. Similarly, Level 3 corresponds to the median proficiency of students in 29 participating countries and economies. Across OECD countries, on average, 27% of students scored at Level 3, the largest share amongst the seven proficiency levels described in PISA. Similarly, in 30 countries and economies, the largest share of students performed at Level 3 (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Proficiency at Level 4

At Level 4, students can use more sophisticated content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context (the ability to conduct experiments was not assessed in paper-based tests). They can justify an experimental design, drawing on elements of procedural and epistemic knowledge. Level 4 students can interpret data drawn from a moderately complex data set or less familiar contexts and draw appropriate conclusions that go beyond the data and provide justifications for their choices.

On average across OECD countries, 25% of students performed at Level 4 or above, and scored higher than 559 points on the PISA science scale. The largest share of students in B-S-J-Z (China) and Singapore performed at this level (the modal level); Level 4 was also the median level of performance in B-S-J-Z (China) and Singapore (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Proficiency at Level 5

At Level 5, students can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes. They can apply more sophisticated epistemic knowledge to evaluate alternative experimental designs, justify their choices and use theoretical knowledge to interpret information or make predictions. Students at this level can evaluate ways of exploring a given question scientifically and identify limitations in the interpretation of data sets, including sources and the effects of uncertainty in scientific data.

Level 5 on the science scale marks another qualitative difference. Students who can complete Level 5 tasks can be said to be top performers in science in that they are sufficiently skilled in and knowledgeable about science to be able to creatively and autonomously apply their knowledge and skills to a wide variety of situations, including unfamiliar ones.

On average across OECD countries, 6.8% of students were top performers, meaning that they were proficient at Level 5 or 6. Almost one in three (32%) students in B-S-J-Z (China), and more than one in five students in Singapore (21%) performed at this level. In 9 countries/economies (Macao [China], Japan, Finland, Estonia, Korea, Chinese Taipei, Canada, New Zealand and the Netherlands, in descending order of the share of students), between 10% and 14% of all students performed at Level 5 or above. By contrast, in 27 countries/economies, including Colombia (0.5%) and Mexico (0.3%), fewer than one in 100 students was a top performer (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Countries and economies with similar mean performance may have significantly different shares of students who are able to perform at the highest levels in PISA. This is true, for example, in Hong Kong (China) (with a mean performance of 517 points and where 7.8% of students were top performers) and Chinese Taipei (with a mean performance of 516 points and where 11.7% of students were top performers). The smaller share of top-performing students in Hong Kong (China) compared to Chinese Taipei reflects a more narrow variation of student performance around the mean.

Proficiency at Level 6

Students at Level 6 on the PISA science scale can successfully complete the most difficult items in the PISA science assessment. At Level 6, students can draw on a range of inter-related scientific ideas and concepts from the physical, life, and earth and space sciences. They can use procedural and epistemic knowledge to offer explanatory hypotheses of novel scientific phenomena, events and processes that require multiple steps or to make predictions. In interpreting data and evidence, they can discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory, and those based on other considerations. Students at Level 6 can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.

On average across OECD countries, 0.8% of students (or about 1 in 120 students) attained Level 6. B-S-J-Z (China) had the largest proportion of students (7.3%) who scored at this level in science, followed by Singapore (3.8%). In 14 participating countries and economies, between 1% and 2% of students scored at this level, while in the remaining countries/economies, fewer than 1 in 100 students scored at the highest level (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Proficiency below Level 2

The PISA science assessment identified two proficiency levels below Level 2. Students who scored at or below these levels are considered low achievers in science.

Proficiency at Level 1a

At Level 1a, students can use common content and procedural knowledge to recognise or identify explanations of simple scientific phenomena. With support, they can undertake structured scientific enquiries with no more than two variables (the ability to undertake scientific enquiry was not assessed in the paper-based test of science). They can identify simple causal or correlational relationships, and interpret graphical and visual data that require a low level of cognitive ability. Students at Level 1a can select the best scientific explanation for given data in familiar personal, local and global contexts.

On average across OECD countries, 16.1% of students performed at Level 1a and only 5.9% of students performed below Level 1a. In the Dominican Republic, fewer than one in two students (about 47%) attained Level 1a (or a higher level of performance). In 15 countries and economies (including some countries that used the paper-based test of science), the median proficiency level of the 15-year-old student population was within Level 1a (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

Proficiency at Level 1b

At Level 1b, students can use common content knowledge to recognise aspects of simple scientific phenomena. They can identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.[2]

Across OECD countries, 5.2% of students performed at Level 1b and 0.7% performed below Level 1b. In 44 countries and economies, less than 10% of students performed at or below Level 1b (Figure I.7.1, Figure I.7.2 and Table I.B1.3).

No item in the PISA assessment can indicate what students who perform below Level 1b can do. Students who scored below Level 1b may have acquired some elements of science knowledge and skills, but based on the tasks included in the PISA test, their ability can only be described in terms of what they cannot do – and they are unlikely to be able to solve, other than by guessing, any of the PISA tasks. In some countries, more than 1 in 20 students performed below Level 1b: 14% in the Dominican Republic, 10% in Panama, and between 9% and 5% in Lebanon, the Philippines, Georgia and Qatar (in descending order of that share).

Notes

1. Six of the seven levels are aligned with the levels used in describing the outcomes of PISA 2006 (ranging from the highest, Level 6, to Level 1a, formerly known as Level 1). These levels and their respective descriptors are still applicable to the paper-based assessment of science.

2. Descriptions of what students can do at Level 1b are based on items included in the PISA 2015 science assessment. In 2018, only one item in the paper-based test of science was located at this level; the easiest tasks included in the PISA 2018 computer-based test of science were located at Level 1a. It is nevertheless possible to estimate for every student the likelihood of scoring at Level 1b, based on how he or she responded to Level 1a tasks (for a student whose proficiency lay below the level of difficulty of the task, the probability of a correct response varied between 0 and 62%, depending on how far below the task's difficulty the student's proficiency lay) and on the science performance of students with similar response patterns in 2015. From these individual estimates of the (posterior) likelihood of performance, it is also possible to obtain country-level estimates of the share of students at each proficiency level. See the *PISA 2018 Technical Report* (OECD, forthcoming[3]) for details.

References

OECD (2018), "PISA for Development Science Framework", in *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264305274-6-en. [2]

OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education,* PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264266490-en. [1]

OECD (forthcoming), *PISA 2018 Technical Report,* OECD Publishing, Paris. [3]

8

Where did performance change between 2015 and 2018?

This chapter discusses short-term changes in student performance – both in mean performance and in the performance distribution – in the PISA assessment between 2015 and 2018.

This volume has so far focused on performance in reading, mathematics and science as measured by the 2018 round of the PISA assessment. However, PISA allows for more than just a snapshot of an education system's performance at a given moment: as a long-term study, dating back to 2000, PISA gives countries and economies an opportunity to see how their performance has evolved over the course of almost two decades.

Chapter 9 discusses long-term trends in student performance. This chapter examines changes in performance between the previous PISA assessment, which took place in 2015, and the latest 2018 assessment. Any changes in performance over such a short period of time can likely be related to, if not attributed to, changes in education policy, in the learning environment (both in and outside of school), and in the composition of student populations that affected children who were 15 years old between 2015 and 2018 (i.e. those born between 1999 and 2002).

What the data tell us

- On average across OECD countries, mean performance in reading, mathematics and science remained stable between 2015 and 2018.
- There were large differences between individual countries and economies in how their performance changed between 2015 and 2018. For example, mean performance in reading improved in 4 countries and economies (Macao [China], the Republic of North Macedonia, Singapore and Turkey), declined in 13 countries/economies, and remained stable in the remaining 46 countries/economies.
- Between 2015 and 2018, the performance distribution in both reading and mathematics widened, on average across OECD countries; but the performance distribution in science neither widened nor narrowed significantly during that period.

In order to attribute changes in performance across PISA cycles to changes in student learning or to differences in the composition of student populations, the PISA test and how it was conducted would have had to remain equivalent from cycle to cycle. Differences in how the test was conducted – such as the length of the test, whether students take the test on paper or on computer, or whether they sit the test in the morning or the afternoon – could affect student motivation and performance; therefore, these differences must be monitored and minimised.

Overall, PISA 2018 and PISA 2015 were conducted in much the same way:

- As in 2015, the vast majority of students who sat the PISA 2018 assessment answered questions in just two subjects, devoting one hour to each: the major domain (reading in 2018, science in 2015) and one other domain (OECD, forthcoming[1]). In previous rounds of PISA, the number of subjects varied more across students: while large numbers of students were tested in two subjects, a significant proportion was tested in three subjects within the same two-hour testing period.
- The assessment was primarily conducted on computer in both 2015 and 2018, whereas it was conducted on paper in 2012 and earlier. While measures were taken in 2015 to align the computer-based tests with the original paper-based scales, these measures were implemented mainly at the international level. Country-specific differences in familiarity with computers, or in student motivation when taking the test on computer or on paper, could still interfere with performance trends (OECD, 2016[2]). For most countries, this mode-related source of uncertainty no longer existed when comparing 2015 and 2018. Furthermore, test administration was more regimented when computer-based assessments were used as there was less room to deviate from standard procedures (e.g. when to take breaks and how long such breaks last).

Annex A8 further explores differences in students' effort and motivation across countries and over time.

At the same time, it is important to assess the impact of using different test items in different years,[1] resulting in performance scales that are not identical. This potential source of error when examining changes in PISA results is summarised by the **link error**, which provides an estimate of the shift in the same subject scale used in two different years. The reporting scales for a subject between two different years are uniformly misaligned by a certain amount. The magnitude of this amount and its direction (i.e. whether a score in one year is equivalent to a higher or lower score in the other year) is unknown, but it is on the order of the size of the link error. For example, the link error between 2015 and 2018 in the reading assessment is roughly four score points, and hence a change of up to eight score points in a country's mean reading performance between 2015 and 2018 would not be significant as it could easily be attributed to the link error.

The link error between 2015 and 2018, however, is noticeably smaller than the link error between other PISA cycles (e.g. between 2012 and 2018, or between 2012 and 2015). In addition to the two reasons listed above concerning how PISA was conducted in 2015 and 2018, there are two further reasons for this smaller link error:

- There were more items in common between the 2015 and 2018 assessments than there were between previous sets of assessments.
  - The 2015 and 2018 mathematics assessments were virtually identical, as they were both minor domains based on the 2012 PISA mathematics framework.
  - The items in the 2018 science assessment were a subset of the items in the 2015 science assessment; the majority of these items were created in 2015 to reflect the updated PISA 2015 science framework and hence differ from items used in assessments prior to 2015.
  - Although new items were developed for the PISA 2018 reading assessment to reflect its new framework (see Chapter 1), a large number were retained from the items developed for PISA 2009 and used between PISA 2009 and 2015.
- In contrast to the procedures used in previous cycles, the characteristics of trend items (those items that were also used in previous cycles of PISA; in this case, in 2015) were assumed to be identical in 2015 and 2018. In practice, item characteristics in 2018 were assumed to be identical to those in 2015, unless there was sufficient evidence of non-equivalence. This resulted in more consistent measurement scales across cycles, and reduced the link error.[2] Items that were unique to one year, however, did not aid in linking scales across years.

In summary, changes between 2015 and 2018 were more precisely estimated than changes involving earlier years. This chapter explores these short-term changes in performance.

CHANGES BETWEEN 2015 AND 2018 IN MEAN PERFORMANCE

Figure I.8.1 shows the changes between 2015 and 2018 in mean performance in reading. On average across OECD countries, mean performance in reading did not change significantly during the period. The decline in performance was most pronounced in Georgia and Indonesia, where it exceeded 20 score points; it exceeded 10 score points in Colombia, the Dominican Republic, Japan, Luxembourg, the Netherlands, Norway, the Russian Federation (hereafter "Russia") and Thailand.

By contrast, several countries/economies saw significant improvements in reading performance. The largest were seen in the Republic of North Macedonia (hereafter "North Macedonia") (41 score points) and Turkey (37 score points), while improvements of between 10 and 20 score points were observed in Macao (China) and Singapore (Figure I.8.1).

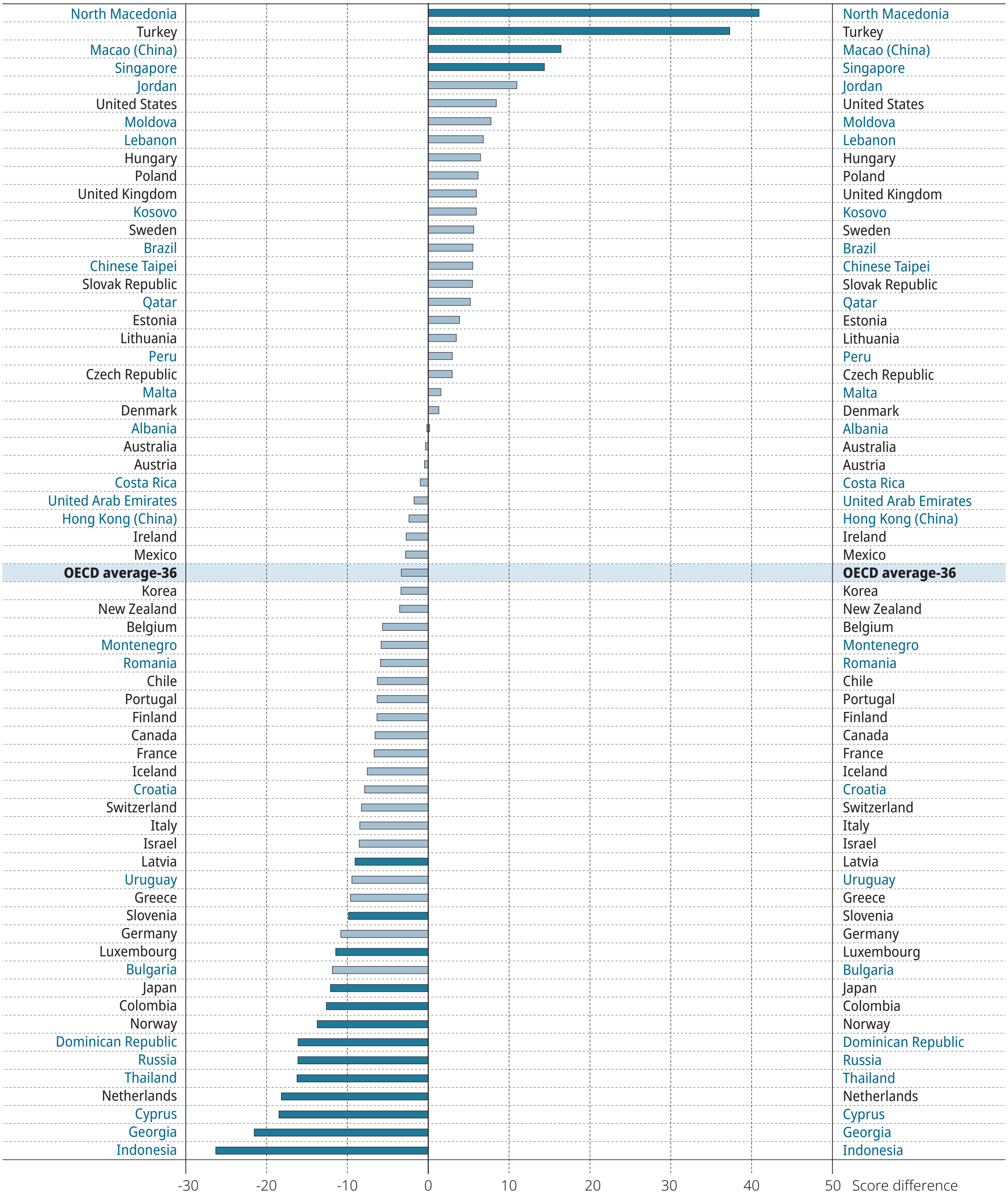
On average across OECD countries, no significant change in either mathematics or science performance was observed between 2015 and 2018. Mathematics performance declined in only three countries/economies (Malta, Romania and Chinese Taipei) during the period, while it improved by over 10 score points in 11 countries/economies (Albania, Jordan, Latvia, Macao [China], Montenegro, North Macedonia, Peru, Poland, Qatar, the Slovak Republic and Turkey). Improvement was notable in Turkey (33 score points), Albania (24 score points) and North Macedonia (23 score points) (Table I.B1.10).

Country-level improvements in science performance were far less common. Improvements of 10 points or more between 2015 and 2018 were observed in only four countries/economies: Turkey (43 score points), North Macedonia (29 score points), Jordan (21 score points) and Macao (China) (15 score points). Science performance declined by at least 10 score points in seven countries/economies: Georgia (28 score points), Bulgaria (22 score points), Chinese Taipei (17 score points), Kosovo (14 score points), Italy (13 score points), Albania (10 score points) and Switzerland (10 score points) (Table I.B1.12).

Most countries and economies did not observe significant changes in performance between 2015 and 2018, when considering each subject independently. This is to be expected. A lack of improvement over three years is not necessarily a cause for concern: education is cumulative and any changes in policy are both incremental and can take years, if not an entire cohort of school-aged children, to have an effect. Moreover, the precision with which differences can be measured means that differences that may be significant in the long term are not deemed significant in the short term. Indeed, in 24 countries and economies out of the 63 that took part in both PISA 2015 and 2018 (Austria, Belgium, Brazil, the Czech Republic, Chile, Costa Rica, Croatia, Estonia, France, Germany, Greece, Hong Kong [China], Hungary, Ireland, Israel, Korea, Lebanon, Lithuania, Mexico, Moldova, New Zealand, Sweden, the United Arab Emirates and the United States), no significant change in performance was observed, between 2015 and 2018, in any of the three core subjects that PISA assessed (Table I.8.1).

During the period, performance improved across all three subjects in Macao (China), North Macedonia and Turkey; and performance improved across two subjects and stayed stable in the third in Jordan and Poland (Table I.8.1).

Figure I.8.1 Change between 2015 and 2018 in mean reading performance



Notes: Statistically significant differences between PISA 2015 and PISA 2018 are shown in a darker tone (see Annex A3).

The change in reading performance between 2015 and 2018 for Spain is not reported; see Annex A9. OECD average-36 refers to the arithmetic average across all OECD countries, excluding Spain.

Countries and economies are ranked in descending order of the change in reading performance between PISA 2015 and 2018.

Source: OECD, PISA 2018 Database, Table I.B1.10.

StatLink https://doi.org/10.1787/888934028691

Encouragingly, in no country or economy did performance decline across all three subjects. However, in seven countries/economies – Georgia, Japan, Luxembourg, Malta, Norway, Slovenia and Chinese Taipei – performance declined in two subjects and remained stable in the third (Table I.8.1).

Table I.8.1 **Change between 2015 and 2018 in mean performance in reading, mathematics and science**

	Reading	Mathematics	Science
Mean performance improved between 2015 and 2018	Macao (China), North Macedonia, Singapore, Turkey	Albania, Iceland, Jordan, Latvia, Macao (China), Montenegro, North Macedonia, Peru, Poland, Qatar, the Slovak Republic, Turkey, the United Kingdom	Jordan, Macao (China), North Macedonia, Poland, Turkey
Mean performance did not change significantly between 2015 and 2018	**OECD average-36**, Albania, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, Costa Rica, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong (China), Hungary, Iceland, Ireland, Israel, Italy, Jordan, Korea, Kosovo, Lebanon, Lithuania, Malta, Mexico, Moldova, Montenegro, New Zealand, Peru, Poland, Portugal, Qatar, Romania, the Slovak Republic, Sweden, Switzerland, Chinese Taipei, the United Arab Emirates, the United Kingdom, the United States, Uruguay	**OECD average-37**, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, the Czech Republic, Denmark, the Dominican Republic, Estonia, Finland, France, Georgia, Germany, Greece, Hong Kong (China), Hungary, Indonesia, Ireland, Israel, Italy, Japan, Korea, Kosovo, Lebanon, Lithuania, Luxembourg, Mexico, Moldova, the Netherlands, New Zealand, Norway, Portugal, Russia, Singapore, Slovenia, Spain, Sweden, Switzerland, Thailand, the United Arab Emirates, the United States, Uruguay	**OECD average-37**, Austria, Belgium, Brazil, Chile, Colombia, Costa Rica, Croatia, the Czech Republic, the Dominican Republic, Estonia, France, Germany, Greece, Hong Kong (China), Hungary, Iceland, Indonesia, Ireland, Israel, Korea, Latvia, Lebanon, Lithuania, Mexico, Montenegro, Moldova, the Netherlands, New Zealand, Peru, Qatar, Romania, Russia, Singapore, the Slovak Republic, Sweden, Thailand, the United Arab Emirates, the United Kingdom, the United States
Mean performance declined between 2015 and 2018	Colombia, the Dominican Republic, Georgia, Indonesia, Japan, Latvia, Luxembourg, the Netherlands, Norway, Russia, Slovenia, Thailand	Malta, Romania, Chinese Taipei	Albania, Australia, Bulgaria, Canada, Denmark, Finland, Georgia, Italy, Japan, Kosovo, Luxembourg, Malta, Norway, Portugal, Slovenia, Spain, Switzerland, Chinese Taipei, Uruguay

Notes: The change in reading performance between 2015 and 2018 for Spain is not reported; see Annex A9. OECD average-36 refers to the arithmetic average across all OECD countries, excluding Spain.
Source: OECD, PISA 2018 Database, Tables I.B1.10, I.B1.11 and I.B1.12.

CHANGES BETWEEN 2015 AND 2018 IN THE PERFORMANCE DISTRIBUTION

The stability in mean performance across OECD countries and in most PISA-participating education systems masks changes in the distribution of student performance. One way this can be seen is by examining the percentiles of student performance. The 10th percentile is the point on the scale below which 10% of students score. In other words, if all students were ranked from lowest- to highest-scoring, the 10th percentile would be the highest-scoring of the lowest-performing 10% of students. Likewise, the 90th percentile is the point on the scale below which 90% of students score (or, conversely, above which only 10% of students score). The median, or 50th percentile, divides the performance distribution into two equal halves, one above and one below that position on the scale.

The subject whose scales should be most comparable between 2015 and 2018 is mathematics, as the assessment items were virtually identical.[3] No significant change was observed in any of the percentiles of the performance distribution between the 10th and 90th, on average across OECD countries, indicating that neither the strongest- nor the weakest-performing students saw an improvement or a decline in performance between 2015 and 2018. However, the inter-decile range (the gap between the 10th and 90th percentiles, and a measure of the dispersion of student performance) increased by 4 score points between 2015 and 2018, on average across OECD countries (Tables I.B1.14 and I.B1.29). This is possible because, although changes in percentiles across time are affected by the offset between scales in different years (i.e. the link error), which can render their measurements less precise, the inter-decile range is not affected by this offset and is thus measured with greater precision.

There was no significant narrowing or widening in the dispersion of the performance distribution in science between 2015 and 2018, on average across OECD countries. There was also no significant change in performance amongst either the strongest- or the weakest-performing students (Tables I.B1.15 and I.B1.30).

Results from PISA 2015 and 2018 indicated that, on average across OECD countries, the score-point difference in reading performance between weaker and stronger students increased by 11 points during that period. The lower level of precision in measuring changes in performance over time, however, made it impossible to state with confidence that stronger students saw an improvement in their performance, or that weaker students saw a decline in theirs (Tables I.B1.13 and I.B1.28).[4, 5]

The discussion above only applies to the average trend across OECD countries; the distribution in performance in individual countries and economies has evolved differently. For example, the inter-decile range in mathematics performance widened significantly in 8 countries and economies (as did the OECD average), while it narrowed significantly in 2 countries/economies and did not change significantly in the remaining 53 countries/economies for which comparable data for 2015 and 2018 were available (Table I.8.2).

Moreover, there were various reasons for why the inter-decile range changed (or did not change) in these countries/economies. For example, the following could explain why the inter-decile range widened between 2015 and 2018:

- Weaker students became weaker and stronger students became stronger.
- Weaker students became weaker but no significant change was observed amongst stronger students.
- Stronger students became stronger but no significant change was observed amongst weaker students.
- All students across the distribution became weaker, but weaker students showed a greater decline in performance than stronger students did.
- All students across the distribution became stronger, but stronger students showed greater improvement in performance than weaker students did.
- There were no significant changes observed at individual percentages (i.e. no significant changes observed amongst either stronger or weaker students) but the overall distribution grew wider.[6]

Table I.8.2 lists countries and economies by whether their performance distributions in reading, mathematics and science narrowed, widened or did not change significantly in dispersion (as measured by the inter-decile range). It also shows whether the change, or lack thereof, was primarily due to changes amongst weaker students, stronger students or both (or in the case of a lack of change, neither). For example, stronger students became stronger but there was no significant change in the performance of weaker students in the United Arab Emirates in mathematics (Table I.B1.13).

The only country where the performance distribution widened between 2015 and 2018 in all three subjects was the United Arab Emirates; it widened in two subjects and remained stable in the third in Canada, Germany, Hong Kong (China) and Romania.[7] There was no country in which the performance distribution narrowed in all three subjects, although it narrowed in two subjects and remained stable in the third in Bulgaria, France, Georgia, Malta and Montenegro (Table I.8.2).

Table I.8.2 [1/2] **Change between 2015 and 2018 in the performance distribution in reading, mathematics and science**

	Reading	Mathematics	Science
Widening of the distribution			
Weaker students became weaker; stronger students became stronger	Hong Kong (China)		
Weaker students became weaker; no significant change amongst stronger students	Canada, Finland, Germany, Iceland, Israel, Latvia, Norway	Germany, Luxembourg, Romania	Romania, the United Arab Emirates
Stronger students became stronger; no significant change amongst weaker students	Australia, Estonia, Macao (China), Poland, Singapore, Sweden, Chinese Taipei, United Arab Emirates, United States	United Arab Emirates	
Almost all students became weaker, but weaker students declined more so than stronger students did	Netherlands, Russia		
Almost all students became stronger, but stronger students improved more than weaker students did	Turkey		North Macedonia
No significant change at individual points along the distribution, although overall widening of the dispersion	**OECD average-36**, Denmark, Ireland, Mexico, Switzerland	**OECD average-37**, Canada, Costa Rica, Norway, Thailand	Hong Kong (China), Qatar

...

Table I.8.2 [2/2] **Change between 2015 and 2018 in the performance distribution in reading, mathematics and science**

	Reading	Mathematics	Science
No change in the dispersion of the distribution			
No significant change along most individual points of the distribution	Austria, Belgium, Brazil, Chile, Colombia, Costa Rica, Croatia, Czech Republic, Greece, Hungary, Italy, Korea, Lebanon, Lithuania, Malta, Moldova, New Zealand, Peru, Portugal, Qatar, Romania, Slovak Republic, Slovenia, United Kingdom, Uruguay	Australia, Austria, Belgium, Brazil, Bulgaria, Chile, Colombia, Croatia, Denmark, Dominican Republic, Estonia, Finland, France, Georgia, Greece, Hong Kong (China), Hungary, Iceland, Indonesia, Ireland, Israel, Italy, Japan, Korea, Kosovo, Lebanon, Lithuania, Mexico, Moldova, Netherlands, New Zealand, Portugal, Russia, Singapore, Slovenia, Spain, Sweden, Switzerland, Uruguay	**OECD average-37**, Austria, Belgium, Brazil, Chile, Colombia, Costa Rica, Croatia, Czech Republic, Dominican Republic, Estonia, Finland, Germany, Hungary, Iceland, Indonesia, Ireland, Israel, Japan, Korea, Latvia, Lebanon, Lithuania, Mexico, Moldova, Netherlands, New Zealand, Norway, Peru, Russia, Slovak Republic, Sweden, Switzerland, Thailand, United Kingdom, United States
Stronger students became weaker; no significant change amongst weaker students			Luxembourg, Portugal
Stronger students became stronger; no significant change amongst weaker students		Czech Republic, United Kingdom, United States	
Most students became weaker	Dominican Republic, Indonesia, Japan, Luxembourg, Thailand	Chinese Taipei	Albania, Australia, Canada, Denmark, Italy, Spain, Chinese Taipei, Uruguay
Most students became stronger	North Macedonia	Albania, Jordan, Latvia, Macao (China), North Macedonia, Peru, Poland, Qatar, Slovak Republic, Turkey	Jordan, Macao (China), Montenegro, Poland, Turkey
Narrowing of the distribution			
Weaker students became stronger; stronger students became weaker	Albania		
Stronger students became weaker; no significant change amongst weaker students	Bulgaria, France, Montenegro	Malta	France, Greece, Malta, Singapore, Slovenia
Weaker students became stronger; no significant change amongst stronger students	Jordan, Kosovo		
Almost all students became weaker, but stronger students declined by more than weaker students did	Georgia		Bulgaria, Georgia, Kosovo
Almost all students became stronger, but weaker students improved by more than stronger students did		Montenegro	

Notes: The change in reading performance between 2015 and 2018 for Spain is not reported; see annex A9. OECD average-36 refers to the arithmetic average across all OECD countries, excluding Spain.

Changes in the dispersion of the distribution – widening, narrowing or no change – are measured by the inter-decile range, or the difference in score points between the 90th percentile and the 10th percentile of the student-performance distribution.

Changes in the location of individual percentiles between 2015 and 2018 are estimated with less precision than changes in the mean. For some countries/economies, a significant change in mean performance was observed during the period even though changes in points along the distribution could not be deemed significant.

It is also possible that there was no significant change in the dispersion of the distribution, but that one of the extremities (i.e. the 10th or the 90th percentile) changed significantly, while the other did not. It should be kept in mind that the difference between significant and non-significant changes is, itself, often non-significant.

When there was either a widening or a narrowing of the distribution, there was a change amongst weaker students if student performance at either the 10th or 25th percentile improved or declined and that at the other percentile moved in the same direction or did not change significantly. Likewise, there was a change amongst stronger students if student performance at either the 75th or 90th percentile improved or declined and if that at the other percentile moved in the same direction or did not change significantly. In order to classify a country/economy as one where almost all students became weaker or stronger, at least four of the percentiles examined (the 10th, 25th, 50th, 75th and 90th percentiles) must have declined or improved.

When there was no change in the dispersion of the distribution, at least three individual points along the distribution that were examined (the 10th, 25th, 50th, 75th and 90th percentiles) must have declined or improved in order to say that most students became weaker or stronger in that country/economy. When there was no change in the dispersion of the distribution, student performance at both the 10th and 25th percentiles had to move in the same direction in order to say that weaker students became stronger or weaker; likewise performance at both the 75th and 90th percentiles had to move in the same direction in order to say that stronger students became stronger or weaker.

Source: OECD, PISA 2018 Database, Tables I.B1.13, I.B1.14 and I.B1.15.

Box I.8.1. **Reading trends and changes in the reading framework**

This chapter discusses changes in reading performance between 2015 and 2018 as if they reflected an evolution in students' abilities during the period. This is likely to be the case for changes in mathematics and science performance, as the 2015 and 2018 assessments in these two subjects were identical or a representative subset of one another. But the framework for the reading assessment changed between 2015 and 2018; hence the evolution in reading performance might be attributable to those changes – particularly in students' relative strengths and weaknesses in certain aspects of reading proficiency that were more or less emphasised in 2018 compared to 2015.[8]

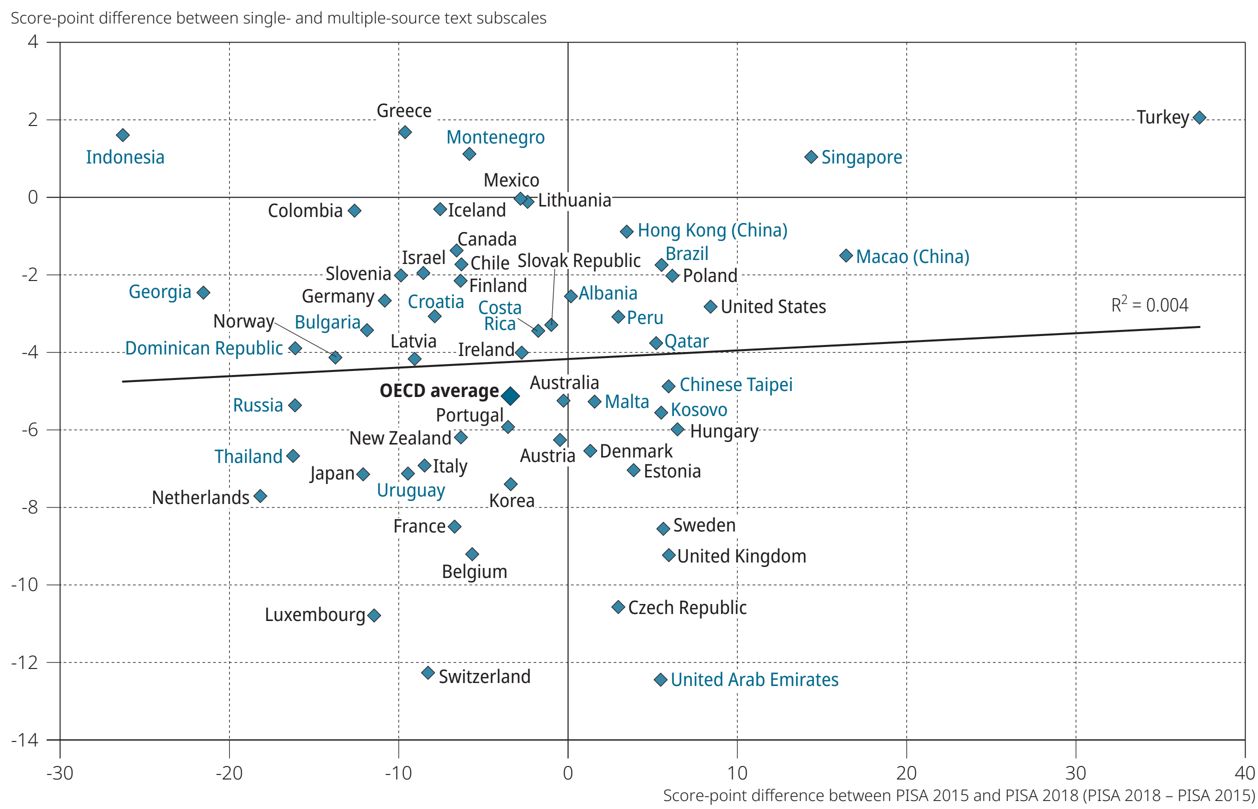
There were two main changes to the framework between 2015 and 2018: the greater focus on multiple-source texts and the inclusion of reading-fluency items. As discussed in Chapter 1, the greater focus on multiple-source texts was made possible by delivering the assessment on computer. Countries and economies whose students were relatively weaker in reading multiple-source texts, for example, might be expected to have more negative trends between 2015 and 2018 than countries whose students were relatively stronger in reading such texts.

Fortunately, it was possible to examine whether this first change to the framework affected student performance. PISA 2018 included a subscale for items that required only a single source and a subscale for those that required reading multiple sources. If changes in performance between 2015 and 2018 were in large part due to the changes in the framework, they would be observed in a correlation between the change in performance and the difference in subscale scores, as both would reflect differences between using single and multiple sources.[9]

Figure I.8.2 shows a scatterplot of the differences in the single- and multiple-source subscales in PISA 2018 versus the change in reading performance between PISA 2015 and PISA 2018. There is no noticeable correlation between the two variables. As a result, it is possible to conclude that the greater emphasis on multiple-source texts in PISA 2018 had a limited impact on changes in reading performance.

Figure I.8.2 **Change in reading performance and score differences on reading subscales**

Change between 2015 and 2018; performance difference in single- and multiple-source reading subscales



Source: OECD, PISA 2018 Database, Tables I.B1.10, I.B1.19 and I.B1.20.
StatLink https://doi.org/10.1787/888934028710

As mentioned above, the other main change in the framework between 2015 and 2018 was the inclusion of reading-fluency items. These items were presented at the beginning of the assessment. They measured whether students could quickly determine whether certain sentences, such as "The red car had a flat tire" or "Airplanes are made of dogs", made sense. These items were used to determine a student's overall reading score but were not part of any subscale. Hence, the part of a student's score that cannot be explained by his or her subscale scores can be taken as a proxy for his or her accuracy in answering reading-fluency items.[10]

There was no correlation between the change in countries/economies' average reading performance between 2015 and 2018 and the estimated accuracy in answering reading-fluency items. Indeed, R^2 values never exceeded 0.04, regardless of how the estimated accuracy was computed or which subscales (reading process or text source) were used, and the (non-significant) direction of the correlation was highly sensitive to the removal of outliers. As with the greater emphasis on multiple-source texts, the inclusion of reading-fluency items does not appear to explain a large degree of the change in reading performance between 2015 and 2018.

However, there seem to be some overarching factors affecting student performance. The cross-country correlation between changes in reading and mathematics performance between 2015 and 2018 is 0.62; that between reading and science is 0.67; and that between mathematics and science is 0.75. Factors that affect performance across subjects seem to play a bigger role in explaining changes in reading performance than either the emphasis on multiple-source texts in, or the addition of reading-fluency items to, the PISA 2018 assessment.

Notes

1. Even the same test items may not retain identical measurement properties across PISA cycles. For example, with the passage of time, respondents may become more familiar with what was initially an unusual item format or component of the test, such as an equation editor or a drawing tool; or they may no longer recognise a particular situation (such as writing postcards or using video recorders) as familiar.

2. Item parameters for items common to 2015 and 2018 were initially constrained to the best-fit values used in 2015. The parameters for 2018 were allowed to vary from the parameters used in 2015 if they poorly fit the PISA 2018 data. Student scores from PISA 2015 were not affected by this procedure, i.e. there was no rescaling of PISA 2015 data.

3. As discussed in note 1 above, there may still be differences in the mathematics scales between 2015 and 2018 even if the same test items were used. However, these are more limited in scope, and have less impact on comparisons between years, than changes in the questions used in the assessment (as occurred in reading and science).

4. In this situation, where there was no significant change in the dispersion of the performance distribution, "weaker students" refer to those at the 10th and 25th percentiles, while "stronger students" refer to those at the 75th and 90th percentiles.

5. Adaptive testing (see Chapter 1) was implemented for the reading assessment in 2018, allowing for greater precision in measuring student performance at both the high and low ends of the performance distribution. Measurement of performance at the low end of the distribution was also enhanced through the addition of reading-fluency items at Levels 1b and 1c. Prior to 2018, the measurement of scores at the extremes was affected by greater uncertainty. Adaptive testing, which presents stronger students with more difficult questions and weaker students with easier questions, and reading-fluency items both improved the precision in measuring these students' scores and therefore the ability to detect significant differences amongst high- or low-achieving students. Results in mathematics and science were not affected by either adaptive testing or the introduction of reading-fluency items.

6. This discussion only considers changes that were statistically significant. As mentioned in the main text, changes in performance over time are subject to link error and therefore measured with less precision than changes in the inter-decile range (i.e. a narrowing or widening of the distribution), which are not subject to link error.

7. The performance distribution in reading widened between 2015 and 2018 in more countries/economies (25) than did the performance distributions in either mathematics (8) or science (5). However, the changes observed in reading performance between 2015 and 2018 may also reflect changes in the framework and design of the test, and must therefore be interpreted with caution.

8. This annex is only concerned with countries that delivered the assessment on computer; the paper-based assessment continued to use the same framework as that used between 2009 and 2015.

9. Differences between the two subscales do not have a substantive meaning. It is not possible to say that countries that are stronger at reading multiple-source texts than single-source texts if their multiple-source text subscale score is higher than their single-source text subscale score – much as it is not possible to say that countries are stronger at reading than mathematics if their reading score is higher than their mathematics score. However, as these two subscales were scaled together to give the overall reading scale, their differences can be compared across countries. For example, a country whose multiple-source text subscale score is higher than their single-source text subscale score is *relatively* stronger at reading multiple-source texts than a country where the two subscale scores are identical. For more information, see Chapter 5.

10. There were two ways to estimate the part of a student's reading score that could not be determined by his or her subscale scores. In the first method, the overall reading score was linearly regressed over the subscale scores; the part of the overall score that could not be explained by the subscale scores was captured by the residual of the regression. In the second method, a composite overall score was created through a weighted average of the subscores; the weights came from the approximate composition of the reading assessment on either the text source (65% single-source text and 35% multiple-source text) or the reading process (25% "locating information", 45% "understanding", and 30% "evaluating and reflecting"). Chapter 1 of this volume provides more details on the breakdown of the PISA 2018 reading assessment. In this second method, the part of a student's reading score that could not be determined by the student's subscale scores is thus the difference between the reading score and the composite, weighted-average score.

    The process of creating an overall reading score is not a simple linear combination of various subscores and the reading-fluency portion of the assessment, so neither of these methods truly captures students' performance on reading-fluency questions. However, these two methods gave highly correlated estimates of performance in reading fluency (R^2 between 0.86 and 0.88).

References

OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, https://dx.doi.org/ 10.1787/9789264266490-en. [2]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

9

Which countries have improved and which countries have declined in performance over their participation in PISA?

This chapter reviews trends in mean performance, and in performance at the various levels of proficiency measured by PISA, between earlier PISA assessments (prior to 2015) and 2018.

PISA 2018 is the seventh round of the international assessment since the programme was launched in 2000. Every PISA test assesses students' knowledge and skills in reading, mathematics and science; each assessment focuses on one of these subjects and provides a summary assessment of the other two (see "What is PISA?" at the beginning of this volume).

The first full assessment of each subject sets the scale and starting point for future comparisons. For reading, trend comparisons are possible starting from 2000. Mathematics was the major domain for the first time in 2003, and science in 2006. This means that it is not possible to measure the change in mathematics performance between PISA 2000 and PISA 2018, nor the change in science performance between PISA 2000, PISA 2003 and PISA 2018. In all subjects, the most reliable way to establish a trend in students' performance is to compare all available results between the first full assessment of each subject and 2018.

Every third assessment provides an opportunity to revisit what it means to be proficient in a core subject and the kinds of contexts in which proficiency manifests itself. With the 2015 assessment, PISA moved its tests to computers; and by 2018, the reading and science tests had been revised to include digital contexts – such as simulations, in science, or online text formats, in reading – in the assessment (the transition to computer-based assessment will be completed in 2021, with the revision of the mathematics framework and test). Because of the changing nature of the test, PISA long-term trends reflect not only whether students have become better at mastering the reading tasks that proficient readers could successfully complete in 2000, or at solving the kinds of mathematics and science problems that were assessed in 2003 or 2006, they also indicate whether students' skills are keeping pace with the changing nature of reading, mathematics and science in contemporary societies.[1]

What the data tell us

- Seven countries and economies saw improvements, on average, in the reading, mathematics and science performance of their students throughout their participation in PISA: Albania, Colombia, Macao (China), the Republic of Moldova, Peru, Portugal and Qatar. Seven countries/economies saw declining mean performance across all three subjects: Australia, Finland, Iceland, Korea, the Netherlands, New Zealand and the Slovak Republic.
- The average trend in reading performance across OECD countries is hump-shaped: the slowly improving trend observed up to 2012 was followed by a decline over the 2012-2018 period, and in 2018, the average performance across OECD countries that participated in both assessments was close to the average performance observed in 2006. A similar, hump-shaped trajectory of mean performance was observed in science too, while the average trend in mathematics was flat.
- No association between trends in mean performance and trends in performance gaps between high- and low-achieving students was observed in any subject. This means that there has been no obvious trade-off between pursuing excellence and closing performance gaps in education.
- Six countries significantly increased enrolment rates in secondary education over their participation in PISA and maintained or improved their mean reading, mathematics and science performance (Albania, Brazil, Indonesia, Mexico, Turkey and Uruguay). This shows that the quality of education does not have to be sacrificed when increasing access to schooling.

For countries that participated in PISA in multiple years, trends in student performance indicate whether and by how much students' skills in reading, mathematics and science are improving. But due to differences in when countries participated in PISA, not all countries and economies can compare their students' performance across every PISA cycle. To better understand a country's/economy's trajectory and maximise the number of countries in the comparisons, this chapter focuses on estimates of the overall direction of trends in student performance, and of how that direction changed over time.[2]

This chapter reports trends in reading, mathematics and science performance for the 65 countries and economies that participated in PISA 2018 and at least one assessment prior to PISA 2015. These countries can compare performance over a period of six years or more, and typically over three or more assessments (except Panama, whose only previous participation was in 2009). Such trends are also referred to as "long-term trends", in contrast to the short-term changes between PISA 2015 and PISA 2018 described in Chapter 8.[3] The methodology underpinning the analysis of trends in performance in this chapter is detailed in Annex A7.

TRENDS IN MEAN PERFORMANCE

Table I.9.1 presents a synopsis of trends in mean performance in reading, mathematics and science. Countries in the top-left cell have shown significant improvements in mean performance, across all three subjects, over their participation in PISA. Countries in the bottom-right cell have shown significant declines in mean performance across all three subjects. (The period considered may differ depending on the subject and the country, and may influence the overall direction of the trend reported in Table I.9.1. Country-specific summaries, by subject, are presented in Annex D).

Table I.9.1 **Trends in mean performance in reading, mathematics and science**

Based on average three-year trend; only countries and economies that participated in PISA 2018 and at least one assessment prior to PISA 2015 are included

		Improving trend in reading	Non-significant trend in reading	Declining trend in reading
Improving trend in mathematics	Improving trend in science	Albania **(ms)**, Colombia **(rm)**, Macao (China) **(r)**, Moldova **(rms)**, Peru **(ms)**, Portugal, Qatar **(rm)**	Georgia **(rms)**, Malaysia **(rms)**, North Macedonia **(ms)**, Turkey **(r)**	
	Non-significant trend in science	Estonia **(rm)**, Israel **(m)**, Montenegro **(rm)**, Poland, Romania **(rm)**, Russia, Serbia **(rm)**,	Brazil, Bulgaria **(m)**, Italy, Kazakhstan **(rms)**, Malta **(rms)**, Mexico,	
	Declining trend in science			
Non-significant trend in mathematics	Improving trend in science	Singapore **(rms)**		
	Non-significant trend in science	Jordan **(rm)**, Chile **(m)**	Argentina **(m)**, Denmark Indonesia, Japan, Latvia, Luxembourg **(r)**, Norway, Panama **(rms)**, Spain **(r)**, the United Arab Emirates **(rms)**, the United Kingdom **(rm)**, the United States, Uruguay **(r)**	Sweden, Thailand
	Declining trend in science	Germany	Austria, Croatia **(rm)**, Greece, Hong Kong (China), Ireland, Lithuania **(rm)**, Slovenia **(rm)**	Costa Rica **(rms)**
Declining trend in mathematics	Improving trend in science			
	Non-significant trend in science		France, Chinese Taipei **(rm)**	
	Declining trend in science		Belgium, Canada, the Czech Republic, Hungary, Switzerland	Australia, Finland, Iceland, Korea, the Netherlands **(r)**, New Zealand, the Slovak Republic **(r)**

Notes: A dark blue bar indicates improving mean performance across all three domains. Medium blue bars indicate improving mean performance across two (out of three) domains, with no decline in the remaining domain. Light blue bars indicate improving mean performance in one domain, with no decline in the remaining domains. White cells indicate non-significant improvement or decline across all three domains, as well as conflicting trends across domains.

A dark grey bar indicates declining mean performance across all three domains. Medium grey bars indicate declining mean performance across two (out of three) domains, with no improvement in the remaining domain. Light grey bars indicate declining mean performance in one domain, with no improvement in the remaining domain.

Not all countries and economies can compare their students' performance over the same period; for many countries and economies, the period considered also differs depending on the subject (the longest possible period is 2000 to 2018 for reading, 2003 to 2018 for mathematics and 2006 to 2018 for science). The overall direction of trends reported in this table may depend on the period over which trends are computed. Letters in parentheses next to the country's/economy's name signal limitations in the period over which trends are computed, which may affect the comparability of trends with those of other countries/economies:

r: Reading trends are computed over a shorter period than 2000, 2001 or 2002 to 2018. For Spain, the reading trend is computed over the 2000-2015 period.

m: Mathematics trends are computed over a shorter period than 2003 to 2018.

s: Science trends are computed over a shorter period than 2006 to 2018.

Source: OECD, PISA 2018 Database, Tables I.B1.10, I.B1.11 and I.B1.12.

Seven countries and economies saw an improving trend in their students' mean reading, mathematics and science performance throughout their participation in PISA: Albania, Colombia, Macao (China), the Republic of Moldova (hereafter "Moldova"), Peru, Portugal and Qatar.

Twelve countries and economies saw significant improvements in students' mean performance in two of the three subjects over their participation in PISA: in reading and mathematics in Estonia, Israel, Montenegro, Poland, Romania, the Russian Federation (hereafter "Russia") and Serbia; in mathematics and science in Georgia, Malaysia, the Republic of North Macedonia (hereafter "North Macedonia") and Turkey; and in reading and science in Singapore. In these countries and economies, student performance in the other subject did not change significantly over their participation in PISA.

Six more countries and economies saw improvements in their performance in mathematics, but not in reading or in science (Brazil, Bulgaria, Italy, Kazakhstan, Malta and Mexico). Three countries saw improvements in reading performance, but not in mathematics or in science (Chile, Germany and Jordan). In Germany, mean science performance actually declined, but over a shorter period, 2006-2018, than is used to establish the trend in reading and mathematics. These and other notable trends are analysed in greater detail in Annex D.

Seven countries/economies showed declining mean performance across all three subjects over their participation in PISA: Australia, Finland, Iceland, Korea, the Netherlands, New Zealand and the Slovak Republic. Six more countries/economies showed declining mean performance in at least two subjects: in mathematics and science, Belgium, Canada, the Czech Republic, Hungary and Switzerland; and in reading and science, Costa Rica.

Twelve countries and economies saw a decline in student performance in one subject only. Eight of these countries/economies saw declines in performance in science, including Germany, where mean performance improved in reading (see above), and Austria, Croatia, Greece, Hong Kong (China), Ireland, Lithuania and Slovenia; two countries/economies showed a decline in mathematics performance (France and Chinese Taipei); and two countries/economies in reading performance (Sweden and Thailand).

For 13 countries/economies, no significant improving or declining trend could be established in any of the subjects.

Curvilinear trajectories of performance

Several countries that participate in PISA can compare their performance over five or more PISA assessments, i.e. over a period of 12 years or longer. But over such a long period, not all trajectories have been linear. The average trend observed over successive PISA assessments does not capture the extent to which this trend corresponds to a steady, almost linear change, or, for example, to an increasingly positive trend or to an improving, but flattening trend (see Figure I.9.1 for an illustration of different possible trajectories). Even countries with no significant average trend may have seen a temporary slump in performance followed by a recovery, or a temporary improvement, followed by a return to prior levels of performance.

Figure I.9.1 categorises countries and economies into nine groups, depending on the shape of the trajectory of their reading performance (Table I.9.2 and Table I.9.3 provide the corresponding information for mathematics and science).[4] Countries with an average improvement across at least five PISA assessments since PISA 2000, 2003 or 2006 are in the top row; countries with no significant positive or negative trend are in the middle row; and countries with a negative trend are in the bottom row. The column indicates whether the trend observed is a steady trend (middle column), or whether it is an accelerating, flattening or reversing trend.

Macao (China) is the only country/economy with a positive and accelerating trend in all three domains (reading, mathematics and science): student performance in Macao (China) improved over time, and more so in recent PISA cycles than in earlier PISA cycles. In contrast, Korea had a negative and accelerating trend in all three domains: student performance in Korea declined over time, with most of the decline observed over the most recent period.

The average trend in reading performance across OECD countries with valid data in all seven assessment cycles is hump-shaped: the slowly improving trend observed up to 2012 (OECD, 2014, pp. 383-384[1]) was followed by a decline between 2012 and 2018; and in 2018, the average performance across OECD countries that participated in every assessment stood close to the average performance observed in 2000, 2003 and 2006.[5] The average trend in science followed a similar hump-shaped trajectory between 2006 and 2018. The average trend is flat in mathematics (all averages refer to the largest set of OECD countries that can compare their results in all assessment cycles for each subject) (Figure I.9.1, Table I.9.2, Table I.9.3, Tables I.B1.10, I.B1.11 and I.B1.12).

However, several countries/economies were able to move to a more positive trajectory in recent years, after a period of stagnation or decline. Sweden showed an improving trend in all three subjects between 2012 and 2018, reversing earlier declines in mean performance. In addition, a U-shaped trajectory in mean reading performance was observed in Argentina, the Czech Republic, Ireland, Slovenia and Uruguay; a similar trajectory in mean mathematics performance was observed in Denmark, Ireland, Jordan, Lithuania, Slovenia and the United Kingdom; and a U-shaped trajectory in mean science performance was observed in Jordan and Montenegro.[6]

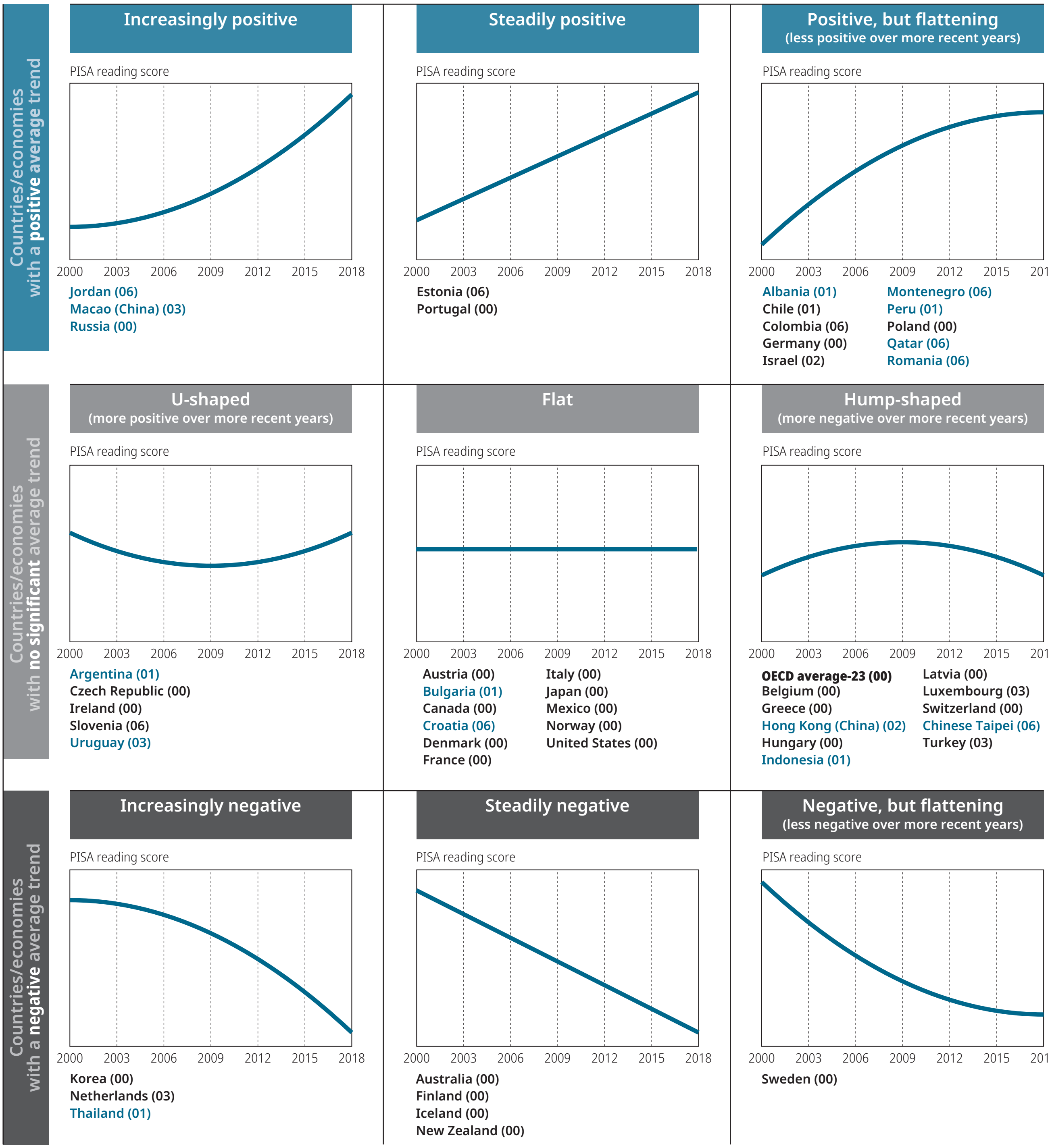
Some countries and economies did not show significant improvement or deterioration over time; their performance remained stable over at least five PISA assessments in each domain. In particular, a "flat" trend was observed in the United States, where reading, mathematics and science scores remained about the same in every PISA assessment, with no particular trend of improvement or decline over its entire participation in PISA.

TRENDS ALONG THE DISTRIBUTION OF PERFORMANCE

Changes in a country's/economy's average performance can result from improvements or declines in performance at different points in the performance distribution. For example, in some countries/economies, improvement is observed along the entire distribution of performance, resulting in fewer students who perform at the lowest levels of proficiency and more students who attain the highest levels of proficiency. In other contexts, average improvement can mostly be attributed to large improvements amongst low-achieving students, with little or no change amongst high-achieving students. This may result in a smaller proportion of low-achieving students, but no increase in the share of top performers.

Figure I.9.1 **Curvilinear trajectories of average performance in reading across PISA assessments**

Direction and trajectory of trend in mean performance



Notes: Figures are for illustrative purposes only. Countries and economies are grouped according to the overall direction of their trend (the sign and significance of the average three-year trend) and to the rate of change in the direction of their trend (the sign and significance of the curvature in the estimate of quadratic trends) (see Annex A7).

Only countries and economies with data from at least five PISA reading assessments are included. Not all countries and economies can compare their students' performance over the same period. For each country/economy, the base year, starting from which reading results can be compared, is indicated in parentheses next to the country's/economy's name ("00" = 2000, "01" = 2001, etc.). Both the overall direction and the change in the direction may be affected by the period considered.

OECD average-23 refers to the average of all OECD countries with valid data in all seven assessments; Austria, Chile, Estonia, Israel, Luxembourg, the Netherlands, the Slovak Republic, Slovenia, Spain, Turkey, the United Kingdom and the United States are not included in this average.

Source: OECD, PISA 2018 Database, Table I.B1.10.

StatLink https://doi.org/10.1787/888934028729

Table I.9.2 **Curvilinear trajectories of average performance in mathematics across PISA assessments**

Direction and trajectory of trends in mean mathematics performance

Countries/economies with a positive average trend	**Increasingly positive**	**Steadily positive**	**Positive, but flattening** (less positive over more recent years)
	Macao (China) (03), Montenegro (06)	Colombia (06), Estonia (06), Poland (03), Russia (03), Turkey (03)	Brazil (03), Bulgaria (06), Israel (06), Italy (03), Mexico (03), Portugal (03), Qatar (06), Romania (06)
Countries/economies with no significant average trend	**U-shaped** (more positive over more recent years)	**Flat**	**Hump-shaped** (more negative over more recent years)
	Denmark (03), Ireland (03), Jordan (06), Lithuania (06), Slovenia (06), Sweden (03), the United Kingdom (06)	**OECD average-29** (03), Austria (03), Croatia (06), Hong Kong (China) (03), Japan (03), Latvia (03), Luxembourg (03), Norway (03), Spain (03), Thailand (03), the United States (03), Uruguay (03)	Chile (06), Germany (03), Greece (03), Indonesia (03)
Countries/economies with a negative average trend	**Increasingly negative**	**Steadily negative**	**Negative, but flattening** (less negative over more recent years)
	Finland (03), Korea (03), Switzerland (03), Chinese Taipei (06)	Australia (03), Canada (03), Hungary (03), the Netherlands (03), New Zealand (03), the Slovak Republic (03)	Belgium (03), the Czech Republic (03), France (03), Iceland (03)

Notes: Countries and economies are grouped according to the overall direction of their trend (the sign and significance of the average three-year trend) and to the rate of change in the direction of their trend (the sign and significance of the curvature in the estimate of quadratic trends) (see Annex A7).

Only countries and economies with data from at least five PISA mathematics assessments are included. Not all countries and economies can compare their students' performance over the same period. For each country/economy, the base year, starting from which mathematics results can be compared, is indicated in parentheses next to the country's/economy's name ("03" = 2003, "06" = 2006). Both the overall direction and the change in the direction may be affected by the period considered.

OECD average-29 refers to the arithmetic mean across all OECD countries, excluding Austria, Chile, Estonia, Israel, Lithuania, Slovenia, the United Kingdom, and excluding Colombia.

Source: OECD, PISA 2018 Database, Table I.B1.11.

Table I.9.3 **Curvilinear trajectories of average performance in science across PISA assessments**

Direction and trajectory of trends in mean science performance

Countries/economies with a positive average trend	**Increasingly positive**	**Steadily positive**	**Positive, but flattening** (less positive over more recent years)
	Macao (China)	Colombia, Turkey	Qatar, Portugal
Countries/economies with no significant average trend	**U-shaped** (more positive over more recent years)	**Flat**	**Hump-shaped** (more negative over more recent years)
	Jordan, Montenegro, Sweden	Chile, France, Indonesia, Israel, Mexico, Russia, Chinese Taipei, the United Kingdom, the United States, Uruguay	**OECD average-36**, Brazil, Bulgaria, Denmark, Estonia, Italy, Japan, Latvia, Luxembourg, Norway, Poland, Romania, Spain, Thailand
Countries/economies with a negative average trend	**Increasingly negative**	**Steadily negative**	**Negative, but flattening** (less negative over more recent years)
	Australia, Germany, Hong Kong (China), Ireland, Korea, Lithuania, Switzerland	Belgium, Canada, Croatia, the Czech Republic, Finland, Greece, Hungary, Iceland, the Netherlands, New Zealand, the Slovak Republic, Slovenia	

Notes: Countries and economies are grouped according to the overall direction of their trend (the sign and significance of the average three-year trend) and to the rate of change in the direction of their trend (the sign and significance of the curvature in the estimate of quadratic trends) (see Annex A7).

Only countries and economies with data from all five PISA science assessments are included. For all countries and economies included in this table, the base year for trends in science performance is 2006.

OECD average-36 refers to the arithmetic mean across all OECD countries (and Colombia), excluding Austria.

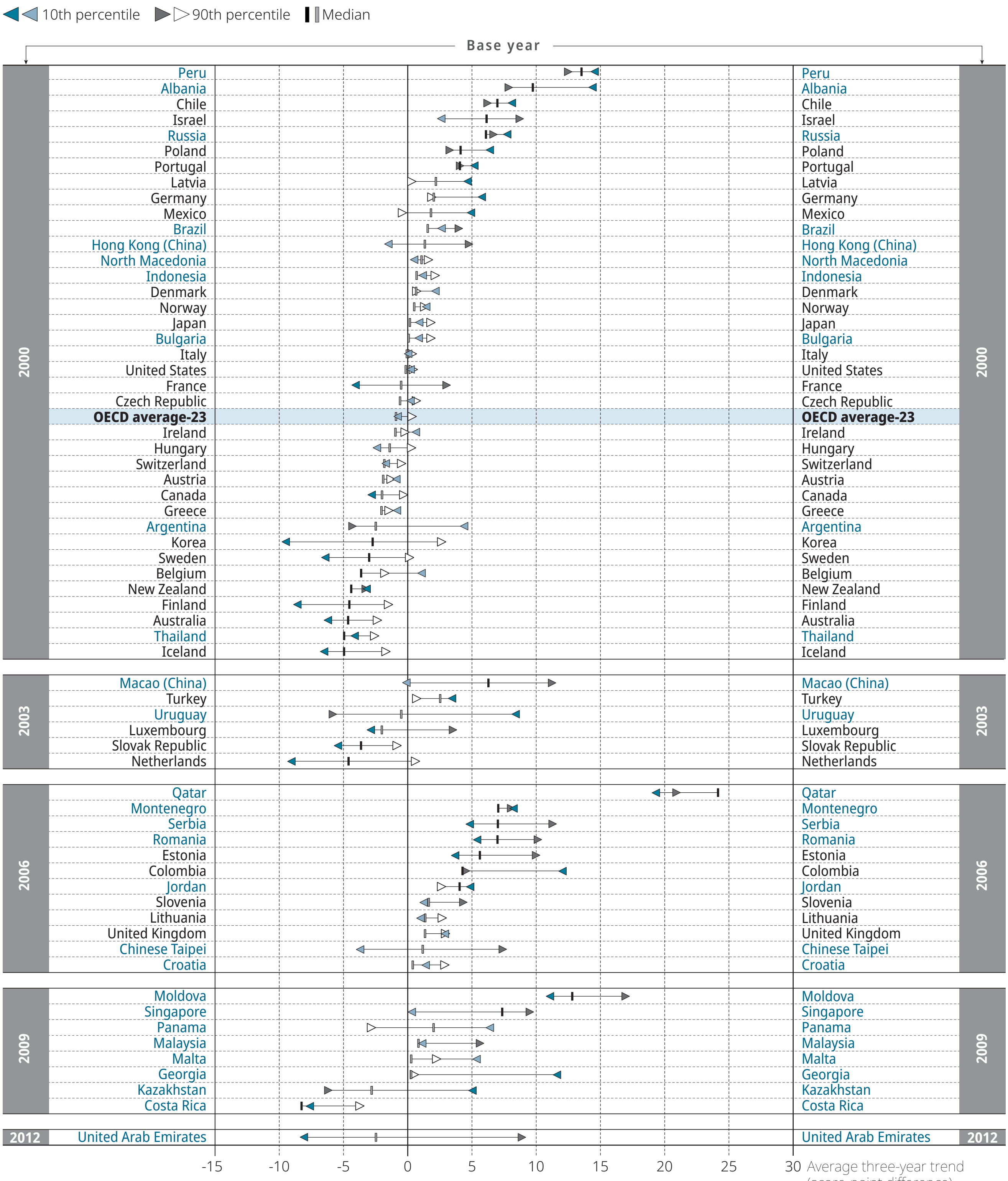
Source: OECD, PISA 2018 Database, Table I.B1.12.

Chapter 4 (Figure I.4.1) shows how performance gaps tend to be greater amongst higher-performing countries/economies, but with many exceptions to this general pattern. Does this cross-country relationship, observed in just one year, imply that there is a trade-off between pursuing excellence and closing performance gaps in education (Parker et al., 2018[2])?

A comparison of trends at the high and low ends of the performance distribution with trends in mean performance suggests that there is no trade-off, in general, between pursuing excellence and reducing learning gaps. Figure I.9.2 shows the linear trend in median performance alongside the trends observed at the 90th and 10th percentiles of the performance distribution (the median performance corresponds to the 50th percentile, or the mid-point, of the performance distribution). Trends at the 10th percentile indicate whether the lowest-achieving 10% of students in a country/economy moved up the PISA scale over time. Similarly, trends at the 90th percentile indicate improvements amongst a country's/economy's high-achieving students (the 90th percentile is the point on the PISA scale below which exactly 90% of students can be found).

Figure I.9.2 **Average three-year trend at different points in the reading proficiency distribution**

Trends in median performance and at the top (90th percentile) and bottom (10th percentile) of the performance distribution



Notes: Values that are statistically significantly different from 0 are marked in a darker tone (see Annex A3).

OECD average-23 refers to the average of all OECD countries with valid data in all seven assessments; Austria, Chile, Estonia, Israel, Luxembourg, the Netherlands, the Slovak Republic, Slovenia, Spain, Turkey, the United Kingdom and the United States are not included in this average.

Countries and economies are grouped by the first PISA assessment starting from which trend comparisons are possible (base year), and are ranked, within each group, in descending order of the average three-year trend in median performance in reading.

Source: OECD, PISA 2018 Database, Table I.B1.13.

StatLink https://doi.org/10.1787/888934028748

Amongst countries and economies with positive trends in reading performance, Albania saw improvements in student performance across the distribution, but more rapidly amongst its lowest-achieving students than amongst its highest-achieving students (Table I.B1.13). As a result, the achievement gap, measured by the inter-decile range (the distance between the 10th and 90th percentiles) shrank by more than 50 points between PISA 2000 and PISA 2018 (Table I.B1.28). Significant reductions in learning gaps were also observed in mathematics and science (for these subjects, trends could only be measured since Albania's second participation in PISA in 2009). Singapore showed a pattern of widening gaps in reading performance since its first participation in 2009, with no improvement in reading at the 10th percentile and increasingly large improvements at the higher percentiles. But in mathematics and science, the gap between the highest- and lowest-performing students in Singapore narrowed over the period, thanks to stronger improvements amongst the lowest-performing students (there was no overall improvement, on average, in mathematics). Macao (China) showed a pattern of widening gaps in reading and science performance (since 2003 and 2006, respectively), but not in mathematics, where differences narrowed.

A significant widening of the gap in reading performance between high- and low-achieving students was observed in the United Arab Emirates, for which trend comparisons are only possible starting from 2012 onwards. Although average reading performance has remained stable since the United Arab Emirates's first participation in PISA 2012, this hides significant improvements amongst high-performing students and rapidly declining performance amongst low-achieving students. Similar patterns of widening performance gaps were observed in mathematics and science (Figure I.9.2, Tables I.B1.29 and I.B1.30).

Overall, across all countries for which PISA can measure long trends in reading performance, there was no significant correlation between changes in average reading performance and changes in the gap in performance (Pearson's *rho* = -0.21).[7] There was also no significant association between trends in mean performance and trends in performance gaps between high- and low-achieving students in mathematics (*rho* = 0.14) and science (*rho* = 0.08).[8] This lack of association suggests that there is no obvious trade-off between raising overall performance and narrowing learning gaps in education (Parker et al., 2018[2]).

Demographic shifts, such as increases in the immigrant population, and changes in enrolment rates (i.e. more disadvantaged 15-year-olds are now enrolled in secondary school than were in previous generations) sometimes contributed to widening disparities in performance. To determine the contribution of these changes to the observed performance trends, "adjusted trends" that neutralise the contribution of demographic and enrolment trends on performance trends are computed (see section "Average three-year trend in performance, adjusted for demographic changes" below).

IMPROVEMENTS AT THE DIFFERENT LEVELS OF PROFICIENCY

PISA assesses reading, mathematics and science skills that are required to participate fully in a knowledge-based society that is increasingly reliant on digital technologies. These range from basic skills that can be considered as minimal capabilities required for further learning, for full participation in most of today's institutions and for non-manual work, to the complex skills that only a few students in most countries have mastered, such as being able to understand and communicate complex information and being able to model complex situations mathematically. Trends in the proportion of low- and top-performing students indicate how the mastery of specific skills (as established in the described proficiency scale) has changed over time.[9]

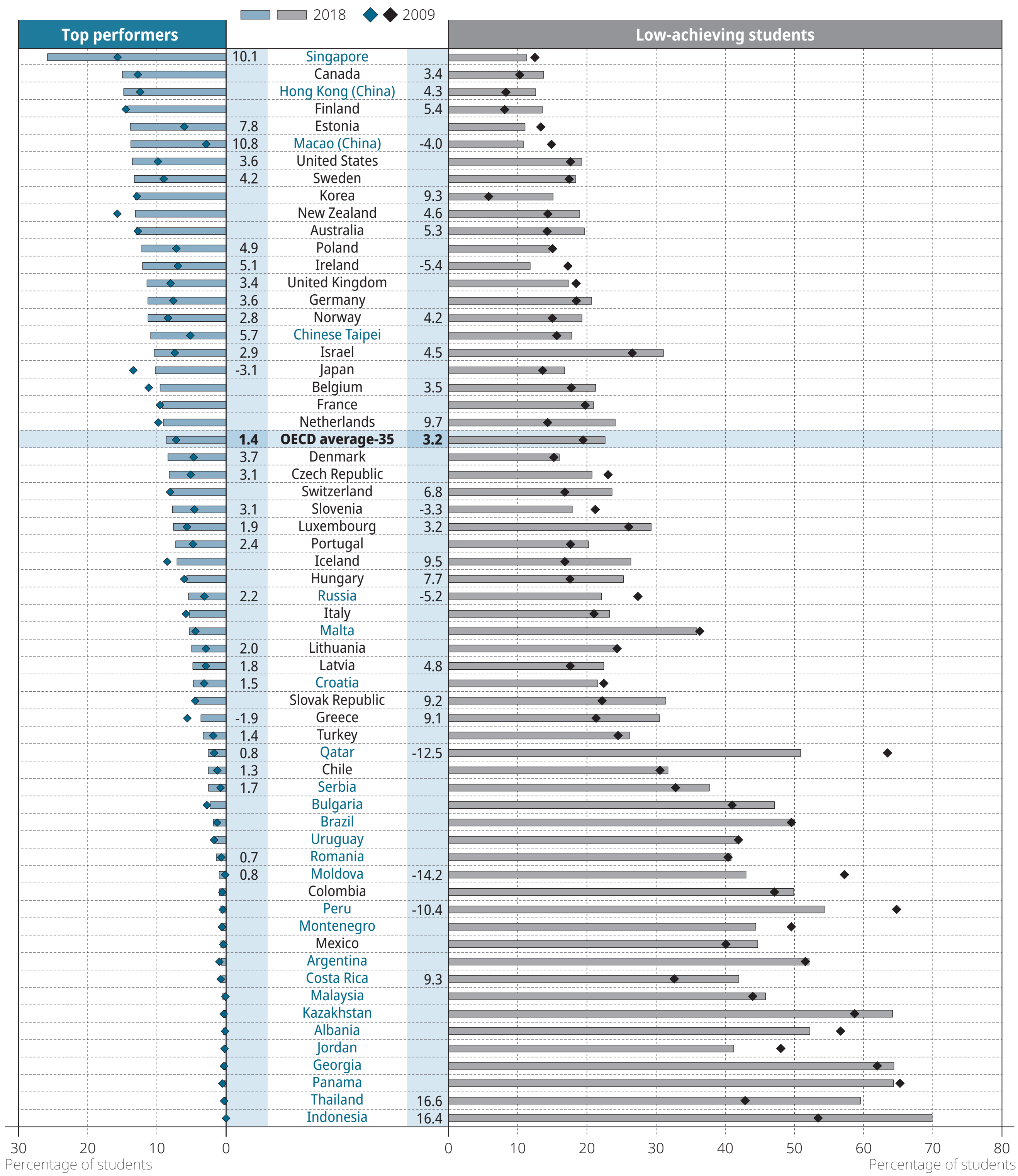
The proportion of students who do not reach Level 2 on the PISA scales (low-achieving students) and the proportion of students who are able to score at Level 5 or 6 (top-performing students) are important indicators of the quality of a country's/economy's talent pool. Trends in the share of low-achieving students indicate the extent to which school systems are advancing towards providing all students with basic literacy and numeracy skills. Trends in the share of top-performing students indicate whether education systems are making progress in ensuring that young people can successfully use their reading, mathematics and science competences to navigate through a volatile, uncertain, complex and ambiguous environment.

On average across OECD countries, the proportion of students scoring below Level 2 in reading increased by 3.2 percentage points between 2009 and 2018, whereas the proportion of students scoring at or above Level 5 increased by 1.4 percentage points (Figure I.9.3). Between 2009 and 2018, seven countries/economies saw reductions in the share of students who scored below Level 2: Ireland, Macao (China), Moldova, Qatar, Russia and Slovenia, which were also able to simultaneously increase the share of students who scored at or above Level 5; and Peru, where the proportion of students scoring at Level 5 or 6 remained stable.

Eighteen more countries/economies saw an increase in the share of students performing at or above Level 5, but no reduction in the share of low-achieving students performing below Level 2. In Israel, Latvia, Luxembourg and Norway, the shares of both low achievers and top performers increased. In Greece, the share of low-achieving students increased, and the share of top-performing students decreased.

Table I.9.4 summarises the information in Figure I.9.3 by grouping countries/economies according to the significance and direction of trends in the share of top-performing and low-achieving students, and presents similar information for mathematics (PISA 2012 to PISA 2018) and science (PISA 2006 to PISA 2018).

Figure I.9.3 **Percentage of low-achieving students and top performers in reading in 2009 and 2018**



Notes: Only countries/economies that participated in both 2009 and 2018 PISA assessments are shown.

Statistically significant changes between PISA 2009 and PISA 2018 in the share of students performing below Level 2 in reading and in the share of students performing at or above Level 5 in reading are shown beside the country/economy name.

OECD average-35 refers to the arithmetic mean across all OECD countries (and Colombia), excluding Austria and Spain.

Countries and economies are ranked in descending order of the percentage of students who scored at or above Level 5 in 2018.

Source: OECD, PISA 2018 Database, Table I.B1.7.

StatLink https://doi.org/10.1787/888934028767

Table I.9.4 **Long-term change in the percentage of low-achieving students and top performers in reading, mathematics and science**

Countries/economies where the ...

... share of low-achieving students (students scoring below Level 2)...	...and the share of top-performing students (students scoring at Level 5 or 6)...	Reading (PISA 2009-PISA 2018)	Mathematics (PISA 2012-PISA 2018)	Science (PISA 2006-PISA 2018)
... decreased ...	... increased	Ireland, Macao (China), Moldova, Qatar, Russia, Slovenia	Albania, Malaysia, Montenegro, Norway, Qatar, Sweden	Colombia, Macao (China), Poland, Portugal, Qatar, Turkey
	... did not change significantly	Peru	Colombia, Jordan, Macao (China), Peru, Slovenia	Brazil, United States
	... decreased			
... did not change significantly ...	... increased	Chile, Croatia, Czech Republic, Denmark, Estonia, Germany, Lithuania, Poland, Portugal, Romania, Serbia, Singapore, Sweden, Chinese Taipei, Turkey, United Arab Emirates, United Kingdom, United States	Kazakhstan, United Arab Emirates	Serbia
	... did not change significantly	Albania, Argentina, Brazil, Bulgaria, Colombia, France, Georgia, Italy, Jordan, Kazakhstan, Malaysia, Malta, Mexico, Montenegro, Panama, Uruguay	Argentina, Austria, Brazil, Bulgaria, Canada, Chile, Costa Rica, Croatia, Czech Republic, Denmark, Estonia, France, Greece, Hungary, Iceland, Indonesia, Israel, Italy, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, Poland, Portugal, Romania, Russia, Serbia, Singapore, Slovak Republic, Spain, Thailand, Turkey, United Kingdom, United States, Uruguay	Argentina, Denmark, Estonia, France, Indonesia, Israel, Japan, Jordan, Korea, Latvia, Lithuania, Mexico, Montenegro, Norway, Romania, Russia, Spain, Sweden, Thailand
	... decreased	Japan	**OECD average-37**, Australia, Belgium, Hong Kong (China), Ireland, Japan, New Zealand, Chinese Taipei	**OECD average-37**, Belgium, Bulgaria, Chile, Czech Republic, Ireland, Italy, Slovenia, United Kingdom, Uruguay
... increased ...	... increased	**OECD average-35**, Israel, Latvia, Luxembourg, Norway		
	... did not change significantly	Australia, Belgium, Canada, Costa Rica, Finland, Hong Kong (China), Hungary, Iceland, Indonesia, Korea, Netherlands, New Zealand, Slovak Republic, Switzerland, Thailand		Germany, Luxembourg, Netherlands
	... decreased	Greece	Finland, Germany, Korea, Switzerland	Australia, Austria, Canada, Croatia, Finland, Greece, Hong Kong (China), Hungary, Iceland, New Zealand, Slovak Republic, Switzerland, Chinese Taipei

Notes: Only countries and economies that participated in PISA 2018 and in the reference assessment (PISA 2009 for reading; PISA 2012 for mathematics; PISA 2006 for science) are included in each column. The change in reading performance between 2009 and 2018 for Spain is not reported; see Annex A9.

OECD average-37 refers to the arithmetic average across all OECD countries and Colombia. OECD average-35 refers to the arithmetic mean across all OECD countries (and Colombia), excluding Austria and Spain.

Source: OECD, PISA 2018 Database, Tables I.B1.7, I.B1.8 and I.B1.9.

AVERAGE THREE-YEAR TREND IN PERFORMANCE, ACCOUNTING FOR CHANGES IN ENROLMENT RATES

In most countries, all boys and girls who were born in 2002 were of the correct age to be eligible to sit the PISA 2018 test. (In countries that tested students in the second part of 2018, a 12-month period spanning the years 2002 and 2003 defined the eligible birthdates.) However, age was not the only criterion for eligibility: 15-year-olds also had to be enrolled in school, in 7th grade or higher, at the time of testing.

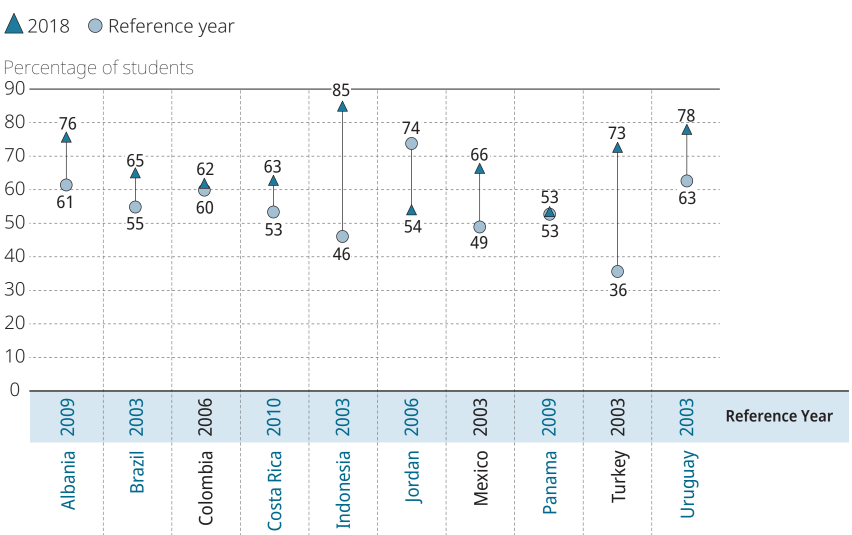
This additional condition might seem redundant in many high-income countries that established universal, free, and sometimes compulsory primary and lower-secondary schooling many decades ago;[10] but because eligibility in PISA is determined by more than just a student's age, the PISA sample does not necessarily represent the entire population of 15-year-olds in many low- and middle-income countries. PISA results thus reflect a combination of 15-year-olds' access to education and the quality of the education that they have received over the course of their lives.

Globally, enrolment in secondary education has expanded dramatically over the past decades. This expansion is also reflected in PISA data, particularly for low- and middle-income countries. Between 2003 and 2018, Indonesia added almost 1.8 million students, and Mexico and Turkey more than 400 000 students to the total population of 15-year-olds eligible to participate in PISA. Brazil and Uruguay, where the total number of 15-year-olds in the country shrank, maintained or increased the number of 15-year-olds eligible to participate in PISA. As a result, PISA coverage – the proportion obtained by dividing the number of PISA-eligible students by the total number of 15-year-olds in a country – increased greatly in all five countries, and most spectacularly in Indonesia (from 46% in 2003 to 85% in 2018) and Turkey (from 36% in 2003 to 73% in 2018). Large increases in coverage, from low initial levels, were also observed in Albania and Costa Rica (since PISA 2009).

Meanwhile, the coverage of 15-year-olds in the PISA sample, which represents students in secondary school, remained stable in Colombia and Panama, and decreased by about 20 percentage points in Jordan. In Jordan, the population of 15-year-olds represented by PISA increased by about 25 000, but the total population of 15-year-olds increased by about 90 000, largely as a result of the massive influx of refugees from neighbouring countries. Refugee children may be enrolled outside of Jordan's formal education system.

Figure I.9.4 **Change in the percentage of 15-year-olds covered by PISA**

Selected countries; 2003 or earliest available year to 2018



Note: Only countries that participated in 2018 and at least one assessment prior to 2015, and where Coverage Index 3 was below 66.6% in their first or latest participation in PISA, are included in the figure.

Countries and economies are listed in alphabetical order.

Source: OECD PISA 2018 Database. Table I.A2.2.

StatLink https://doi.org/10.1787/888934028786

Several factors contributed to lowering the social, economic or institutional barriers that had kept a large proportion of 15-year-olds out of school. Some countries, such as Brazil and Turkey, raised the age at which students can leave compulsory education to over 15; many countries also introduced or strengthened support for at-risk families (e.g. in the form of conditional or unconditional cash transfers). The rapid changes in the economy and the increased urbanisation observed in these countries may also have played a role (UNESCO, 2015[3]).

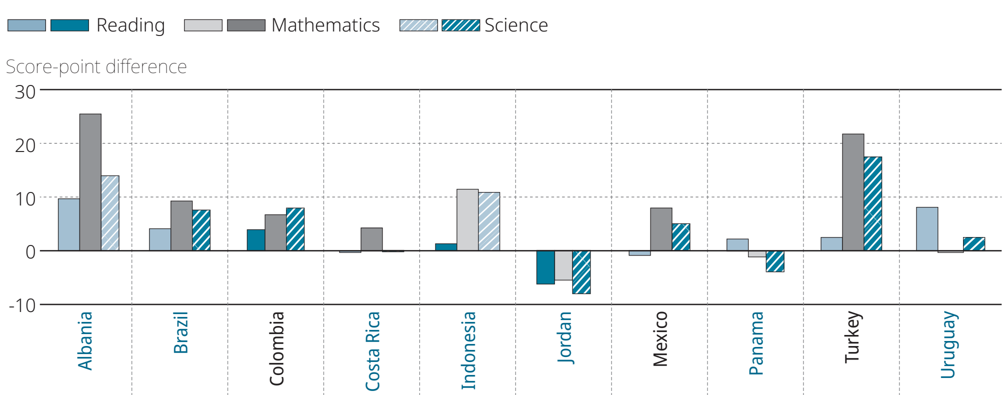
This welcome expansion in education opportunities makes it more difficult to interpret how mean scores in PISA have changed over time. Indeed, increases in coverage can lead to an underestimation of the real improvements that education systems have achieved. Household surveys often show that children from poor households, ethnic minorities or rural areas face a greater risk of not attending or completing lower secondary education (UNESCO, 2015[3]). Typically, as populations that had previously been excluded gain access to higher levels of schooling, a larger proportion of low-performing students will be included in PISA samples (Avvisati, 2017[4]).

The experience of most of the countries shown in Figure I.9.4, however, demonstrates that increases in access to schooling have not, in general, come at the expense of the average quality of education that 15-year-olds receive. In fact, Albania saw significant improvements in its students' average performance in PISA in all three core PISA subjects – reading, mathematics and science – between 2009 and 2018. Turkey saw improvements in its students' average performance in mathematics (between 2003 and 2018) and science (between 2006 and 2018). Brazil and Mexico observed improvements in their students' mean performance in mathematics between 2003 and 2018. Mean results in Indonesia, Panama and Uruguay remained close to those observed in their first year of participation in PISA. Only Costa Rica saw a significant decline in its average reading and science performance over its participation in PISA (2010-2018) (Table I.9.1, Tables I.B1.10, I.B1.11 and I.B1.12).

Under plausible assumptions, this means that amongst countries that expanded enrolment significantly over their participation in PISA, all but one country (Costa Rica) probably saw significant improvements in the level of reading, mathematics and science proficiency attained by the top quarter of 15-year-olds (Figure I.9.5). By considering a population equal in size to 25% of an age group, made up of only the best-performing students in a country, it is possible to monitor the rate of change in PISA performance for a sample of 15-year-olds that was minimally affected by changes in coverage rates over a given period.[11] This analysis shows that the minimum scores in mathematics observed amongst this 25% of top-performing youth increased rapidly (by more than 20 points per three-year period) in Albania and Turkey, and by about 10 points per three-year period in Brazil, Indonesia and Mexico. This suggests that when more disadvantaged children gain access to education for the first time, the remaining students can also benefit.

Figure I.9.5 **Linear trend in the minimum score attained by at least 25% of 15-year-olds**

Selected countries, assuming that in these countries, 15-year-olds not covered by PISA would have performed amongst the bottom 75% had they sat the test (2003 or earliest available year to 2018)



Notes: Only countries that participated in 2018 and at least one assessment prior to 2015, and where Coverage Index 3 was below 66.6% in their first or latest participation in PISA, are included in the figure. For countries that participated in 2003, trends in reading and mathematics performance are measured over a different period than trends in science. See also Figure I.9.4 for relevant context.

Values that are statistically significant are marked in a darker tone (see Annex A3).

Countries and economies are listed in alphabetical order.

Source: OECD, PISA 2018 Database, Tables I.B1.13, I.B1.14, I.B1.15, I.B1.34, I.B1.35 and I.B1.36.

StatLink https://doi.org/10.1787/888934028805

AVERAGE THREE-YEAR TREND IN PERFORMANCE, ADJUSTED FOR DEMOGRAPHIC CHANGES

In some countries, the demographics of the student population and of the PISA sample have changed considerably across PISA assessments. It is possible to analyse the impact of changes in the immigrant background, age and gender of the student population in each country and economy by contrasting the (unadjusted) changes in mean performance, reported in previous sections, with those that would have been observed had the overall profile of the student population been the same, throughout the period, as that observed in 2018. Adjusted trends provide an estimate of what the performance trend would have been if past PISA samples had had the same proportion of immigrant students (first- and second-generation) and the same composition by gender and age (defined in three-month increments) as the target population in 2018.

Adjusted trends for all countries are available in Tables I.B1.40-I.B1.48 in Annex B1 (for more details about the methods used to neutralise changes in the demographic composition of student populations, see Annex A7). On average across the 35 OECD countries that can compare their reading performance between 2009 and 2018, if the student population in 2009 had had the same demographic profile as the population in 2018, the average score in reading would have been 489 points (Table I.B1.40). In reality, the average observed score in 2009 was 491 points (Table I.B1.10). Much of the (non-significant) drop in OECD average performance between 2009 and 2018 (4 points, on average) can therefore be related to the changing demographic composition of student populations, and in particular, to the increase in the shares of first-generation immigrant students in countries where these students tend to score below non-immigrant students; see also Chapter 9 in *PISA 2018 Results* (*Volume II): Where All Students Can Succeed* (OECD, 2019[5]), which presents trends amongst immigrants and non-immigrant students separately. Adjusted and non-adjusted changes in reading performance differ by five score points or more in Qatar (where non-adjusted changes indicate stronger improvement than adjusted changes), and in Germany, Luxembourg, Norway, Sweden and Switzerland (where non-adjusted changes indicate stronger declines than adjusted changes, meaning that some of the declines could be attributed to the changing composition of the student population).[12] For these countries, the counterfactual trends are discussed in greater detail in the context of the country-specific summaries presented in Annex D.

Informative as they may be, adjusted trends are merely hypothetical scenarios that help discern the sources of changes in student performance over time. Observed (unadjusted) trends shown in Figure I.9.1 and throughout this chapter summarise the observed overall evolution of student performance. The comparison of observed trends with hypothetical, adjusted trends can highlight some of the challenges that countries and economies face in improving students' and schools' performance.

Notes

1. In 2018, a few countries continued to assess students using pre-2015 pen-and-pencil tests (see Annex A5). Amongst countries discussed in this chapter, this was the case in Argentina, Jordan, the Republic of Moldova, the Republic of North Macedonia and Romania.
2. The overall direction of a trend is estimated by the linear trend. This represents the average change in student performance per three-year interval, observed over the entire period for which data are available. This period may vary, depending on the country and the subject assessed. Because the rate of change is reported over intervals of three years, the linear trend is referred to as the "three-year trend" in this chapter. Since three years corresponds to the typical interval between two PISA assessments, the average three-year trend can be directly compared to, for example, the change observed between PISA 2015 and PISA 2018, which is described in Chapter 8. For countries and economies that have participated in all PISA assessments, the average three-year trend takes into account up to seven points in time (for reading); for those countries that have valid data for fewer assessments, the average three-year trend takes into account only the valid and available information.
3. Not all OECD countries participated in all PISA assessments. When computing average changes and trends in performance, only those countries with valid data to compare across assessments are included in the average. Because multiple comparisons are often possible within a given table (e.g. between PISA 2000 and PISA 2018 results, but also between PISA 2009 and PISA 2018 results), tables showing trends often include multiple averages.
4. Non-linear trend trajectories are estimated using a regression model, by fitting a quadratic function to the five, six or seven mean estimates available, and taking into account the statistical uncertainty associated with each estimate as well as with comparisons over time (see Annex A7). This is a more robust measure of a country's/economy's trajectory in performance than the successive comparison of mean scores across consecutive assessments because it is less sensitive to one-time statistical fluctuations that may alter a country's/economy's mean performance estimate.
5. All of the pairwise comparisons of performance for the OECD average correspond to non-significant differences; however, the trend line, which reduces the statistical uncertainty by combining information from more than two assessments, shows a significant negative curvature.
6. "Hump-shaped" and "U-shaped" trajectories describe countries for which a significant change in the direction of the trend could be established over their participation in PISA, but no overall improvement or decline was found. The actual shape of a "U-shaped" trajectory may be closer to a "V" or "J" shape (or its mirror image); similarly, the actual shape of a "hump-shaped" trajectory may be similar to an inverse-U shape, but also to an "inverse-J" or "inverse-V" shape.
7. These correlations are measured between the longest difference in mean performance observed in PISA and the difference in the inter-decile range over the same time period, across 64 countries/economies. The Spearman rank correlation is very similar (-0.24).
8. The corresponding Spearman rank-correlation coefficients are 0.04 (mathematics) and -0.08 (science).

9. In this section, the proportions of students at Level 5 and above, and below Level 2, are compared across countries over the same period, starting with the most recent assessment, before 2015, in which a particular domain was the focus of the assessment (the reference year is 2009 in reading; 2012 in mathematics; 2006 in science). For reading and science, due to updates to the assessment framework, the specific abilities that define top-performing and low-achieving students differ between the reference year and 2018, but the same cut-off scores on the equated scales were used to define and compare proficiency levels.

10. The *International Covenant on Economic, Social and Cultural Rights* adopted by the United Nations General Assembly on 16 December 1966 recognises the right of everyone to free primary education and commits its parties to work towards introducing free education at secondary and higher levels (United Nations General Assembly, 1966[9]).

11. The interpretation of these trends requires the additional hypothesis that all the 15-year-olds who were excluded from participation in PISA in past cycles (mostly because they were not in secondary school at age 15) would not have scored above the "adjusted 75th percentile" if they had sat the test. In other words, this analysis relies on the hypothesis that, while the skills and ability of the 15-year-olds who were not eligible to participate in PISA may vary, this variation is bounded below the 75th percentile of the distribution of 15-year-olds' performance in the subjects assessed by PISA. In particular, 15-year-olds who were not in school, or were below grade 7, at the time of the PISA test would not have scored amongst the country's top quarter had they sat the PISA test. No assumption is made about how well these 15-year-olds would have scored if they had received the additional schooling that would have made them eligible to sit the PISA test. If some of the non-eligible 15-year-olds had had greater skills than assumed in this analysis, the 75th percentile estimates on which this analysis is based are, in reality, lower bounds on the *true* 75th percentiles. As the selectivity of PISA samples is attenuated (i.e. Coverage Index 3 increases), the lower bounds can be expected to move closer to the true value. In that context, the reported changes and trends may over-estimate the *true* changes and trends. For a discussion of non-parametric methods for partial identification of trends in the presence of selection, see Blundell et al. (2007[10]).

    It is impossible to know for certain what the PISA score of the 15-year-olds who were not enrolled in school or who were still in grades 1 through 6 would have been, had they been tested. Without attributing an exact score to these students, it is nevertheless possible to assume, with some confidence, that they would have scored in the bottom part of a country's performance distribution (Hanushek and Woessmann, 2008[6]; Spaull and Taylor, 2015[8]; Taylor and Spaull, 2015[7]).

12. In Australia, Portugal and Romania, adjusted and non-adjusted changes in reading performance between PISA 2009 and PISA 2018 also differ by more than five score points. In these countries, however, the major change observed in student demographics is a large increase in the proportion of students whose immigrant status is missing in the data, due to non-response to the questions about country of birth. Adjusted changes must therefore be interpreted with caution because they are based on the assumption that the characteristics of students whose immigrant status is missing are comparable across assessments.

References

Avvisati, F. (2017), "Does the quality of learning outcomes fall when education expands to include more disadvantaged students?", *PISA in Focus*, No. 75, OECD Publishing, Paris, http://dx.doi.org/10.1787/06c8a756-en. [4]

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. [10]

Hanushek, E. and **L. Woessmann** (2008), "The Role of Cognitive Skills in Economic Development", *Journal of Economic Literature*, Vol. 46/3, pp. 607-668, http://dx.doi.org/10.1257/jel.46.3.607. [6]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [5]

OECD (2014), *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, https://dx.doi.org/10.1787/9789264208780-en. [1]

Parker, P. et al. (2018), "Inequity and Excellence in Academic Performance: Evidence From 27 Countries", *American Educational Research Journal*, Vol. 55/4, pp. 836-858, http://dx.doi.org/10.3102/0002831218760213. [2]

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. [8]

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. [7]

UNESCO (2015), *Education for All 2000-2015: Achievements and Challenges. EFA Global Monitoring Report*, UNESCO, https://unesdoc.unesco.org/ark:/48223/pf0000232205 (accessed on 4 September 2019). [3]

United Nations General Assembly (1966), *International Covenant on Economic, Social and Cultural Rights*, p. 3, https://www.refworld.org/docid/3ae6b36c0.html (accessed on 17 August 2019). [9]

10

Measuring global education goals: How PISA can help

This chapter describes how PISA helps countries monitor progress towards the internationally agreed targets of quality and equity in education, and how PISA contributes to improving the capacity of countries to develop relevant data.

In September 2015, the world's leaders gathered in New York to set ambitious goals for the future of the global community. The 17 Sustainable Development Goals (SDGs) adopted by the 70th General Assembly of the United Nations in 2015, otherwise known as the Global Goals or the 2030 Agenda for Sustainable Development, are a universal call for action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity. Developed through an inclusive intergovernmental process, the 2030 Agenda integrates the social, environmental and economic pillars of sustainability with peace and security objectives.

The fourth SDG (SDG 4), to be achieved by 2030, is to: "Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". SDG 4 is to be achieved by meeting ten targets, representing the most comprehensive and ambitious agenda for global education ever attempted.

What the data tell us

- The share of 15-year-old students, in grade 7 and above, who reached a minimum level of proficiency in reading (i.e. at least Level 2 on the PISA scale) ranged from close to 90% in Beijing, Shanghai, Jiangsu and Zhejiang (China), Estonia, Macao (China) and Singapore, to less than 10% in Cambodia, Senegal and Zambia, countries that participated in the PISA for Development assessment in 2017.
- In mathematics, the share of 15-year-old students who attained minimum levels of proficiency (Level 2 and above on the PISA scale) varied even more – between 98% in Beijing, Shanghai, Jiangsu and Zhejiang (China) and 2% in Zambia.
- Disparities in above-minimum proficiency related to socio-economic status were found in all countries and tend to be large. On average across OECD countries, there were only about 7 socio-economically disadvantaged students scoring above minimum levels in reading or mathematics for every 10 advantaged students scoring above these levels.

SDG 4 differs from the Millennium Development Goals (MDGs) on education, which preceded the SDGs and were in place between 2000 and 2015, in the following two ways:

- Like all other SDGs, Goal 4 establishes a universal agenda, and does not differentiate between rich and poor countries. Every single country is challenged to achieve the SDGs.
- Goal 4 puts the quality of education and learning outcomes front and centre. Access, participation and enrolment, which were the focus of the MDG agenda, are still important, and the world is still far from providing equitable access to high-quality education for all. But participation in education is not an end in itself; what matters for people and for development are the skills acquired through education. It is mainly the competencies and character qualities that are developed through schooling, rather than the qualifications and credentials gained, that contribute to people's success and resilience in their professional and personal lives, support individual well-being, and strengthen the prosperity of societies.

In sum, Goal 4 requires all countries to monitor the actual learning outcomes of their young people. PISA, which provides measurement tools to this end, has started to improve, expand and enrich its assessment instruments to help countries in this exercise. This chapter describes how PISA is helping countries monitor progress towards the internationally agreed targets of quality and equity in education, and how PISA contributes to improving the capacity of countries to develop relevant data.

MEASURING COUNTRIES' PROGRESS TOWARDS MEETING GLOBAL EDUCATION TARGETS

By including PISA data in the United Nations' global indicator framework (UNESCO Institute for Statistics, 2019[1]; United Nations Statistics Division, 2019[2]), the global community has recognised the role of PISA in monitoring progress towards the SDG for education over the next decade. PISA data are used for monitoring progress in the proportion of children and young people who, at the end of lower secondary education, have achieved at least minimum proficiency in reading and mathematics (SDG global indicator 4.1.1c). PISA-based indicators are also used to measure how close countries are to meeting other targets, particularly those related to equity and education for sustainable development.

In 2018, PISA assessed the reading, mathematics and science performance of 15-year-old students in 79 countries and economies. An additional seven countries collected comparable data about their students' foundational skills in 2017, as part of the PISA for Development initiative.[1] That project enhanced the PISA paper-based tests to provide more nuanced measures of the reading, mathematics and science competences of 15-year-olds who scored at or below proficiency Level 2. These enhanced pen-and-paper tests will be offered to all countries that wish to continue to assess their students in pen-and-paper mode starting with PISA 2021. A pilot assessment of the reading and mathematics skills of 15-year-olds who, for whatever reason, do not attend school, was also conducted in 2018 in five countries (Panama, and four countries that took part in PISA for Development: Guatemala, Honduras, Paraguay and Senegal). Results of that assessment will be released in the first quarter of 2020.

SDG Target 4.1

The global indicator for the first SDG 4 target is a measure of the "Proportion of children and young people [at different stages of their education career] achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex". PISA provides both a way of defining what "minimum proficiency level" means, through its described scale of proficiency, and a way of measuring this proportion, in an internationally comparable manner, amongst students who are close to the end of lower secondary education (or have recently completed lower secondary education). The UNESCO-led Technical Co-operation Group (TCG) on the Indicators for SDG 4 has officially recognised PISA as a source of data for this global indicator (UNESCO Institute for Statistics, 2019[1]).

Table I.10.1 [1/2] **Snapshot of minimum achievement in reading and mathematics**

			Achievement at Level 2 and above in reading			Achievement at Level 2 and above in mathematics		
			Share of 15-year-old students achieving at Level 2 or above (2018)[2]	Change in this share, expressed as a percentage of the PISA target population (2009 to 2018)[3]	Change in this share, expressed as a percentage of all 15-year-olds (2009 to 2018)[3]	Share of 15-year-old students achieving at Level 2 or above (2018)[2]	Change in this share, expressed as a percentage of the PISA target population (2012 to 2018)[4]	Change in this share, expressed as a percentage of all 15-year-olds (2012 to 2018)[4]
		CI3[1]	%	% dif.	% dif.	%	% dif.	% dif.
OECD	Australia	0.89	80.4	-5.4	N.A.	77.6	N.S.	N.A.
	Austria	0.89	76.4	-4.1	N.A.	78.9	N.S.	N.A.
	Belgium	0.94	78.7	-3.5	N.A.	80.3	N.S.	N.A.
	Canada	0.86	86.2	-3.5	N.A.	83.7	N.S.	N.A.
	Chile	0.89	68.3	N.S.	N.A.	48.1	N.S.	N.A.
	Colombia	0.62	50.1	N.S.	N.S.	34.6	8.4	N.S.
	Czech Republic	0.95	79.3	N.S.	N.A.	79.6	N.S.	N.A.
	Denmark	0.88	84.0	N.S.	N.A.	85.4	N.S.	N.A.
	Estonia	0.93	88.9	N.S.	N.A.	89.8	N.S.	N.A.
	Finland	0.96	86.5	-5.4	N.A.	85.0	-2.7	N.A.
	France	0.91	79.1	N.S.	N.A.	78.7	N.S.	N.A.
	Germany	0.99	79.3	N.S.	N.A.	78.9	-3.4	N.A.
	Greece	0.93	69.5	-9.2	N.A.	64.2	N.S.	N.A.
	Hungary	0.90	74.7	-7.7	N.A.	74.4	N.S.	N.A.
	Iceland	0.92	73.6	-9.5	N.A.	79.3	N.S.	N.A.
	Ireland	0.96	88.2	5.4	N.A.	84.3	N.S.	N.A.
	Israel	0.81	68.9	-4.5	N.A.	65.9	N.S.	N.A.
	Italy	0.85	76.7	N.S.	N.A.	76.2	N.S.	N.A.
	Japan	0.91	83.2	N.S.	N.A.	88.5	N.S.	N.A.
	Korea	0.88	84.9	-9.3	N.A.	85.0	-5.9	N.A.
	Latvia	0.89	77.6	-4.9	N.A.	82.7	N.S.	N.A.
	Lithuania	0.90	75.6	N.S.	N.A.	74.4	N.S.	N.A.
	Luxembourg	0.87	70.7	-3.3	N.A.	72.8	N.S.	N.A.
	Mexico	0.66	55.3	N.S.	N.S.	43.8	N.S.	N.S.
	Netherlands	0.91	75.9	-9.8	N.A.	84.2	N.S.	N.A.
	New Zealand	0.89	81.0	-4.6	N.A.	78.2	N.S.	N.A.
	Norway	0.91	80.7	-4.3	N.A.	81.1	3.4	N.A.
	Poland	0.90	85.3	N.S.	N.A.	85.3	N.S.	N.A.
	Portugal	0.87	79.8	N.S.	N.A.	76.7	N.S.	N.A.
	Slovak Republic	0.86	68.6	-9.2	N.A.	74.9	N.S.	N.A.
	Slovenia	0.98	82.1	3.3	N.A.	83.6	3.7	N.A.
	Spain	0.88	M	M	M	75.3	N.S.	N.A.
	Sweden	0.86	81.6	N.S.	N.A.	81.2	8.3	N.A.
	Switzerland	0.89	76.4	-6.8	N.A.	83.2	-4.4	N.A.
	Turkey	0.73	73.9	N.S.	10.9	63.3	N.S.	6.3
	United Kingdom	0.85	82.7	N.S.	N.A.	80.8	N.S.	N.A.
	United States	0.86	80.7	N.S.	N.A.	72.9	N.S.	N.A.
	OECD average-35a	0.88	77.4	-3.2	N.A.	M	M	M
	OECD average-37	0.88	M	M	M	76.0	N.S.	N.A.

1. CI3: Coverage Index 3, corresponding to the proportion of 15-year-olds who are represented by the PISA sample. For Paraguay, Coverage Index 3 is reported as missing; see Chapter 11 in the *PISA for Development Technical Report* (OECD, 2018[3]) (https://www.oecd.org/pisa/pisa-for-development/pisafordevelopment2018technicalreport/PISA_D_Chapter_11_SamplingOutcomes.pdf, accessed on 28 August 2019).
2. Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia: data refer to 2017 and were collected as part of the PISA for Development assessment.
3. Austria, OECD average-37 and United Arab Emirates: 2012 to 2018; Dominican Republic, Kosovo, Lebanon and North Macedonia: 2015 to 2018.
4. Dominican Republic, Georgia, Kosovo, Lebanon, Malta, Moldova and North Macedonia: 2015 to 2018.

N.S.: not significant.
N.A.: not applicable (Coverage Index 3 is above 0.75).
M: missing due to data availability.

Source: OECD, PISA 2018 Database, Table I.B1.49.

StatLink https://doi.org/10.1787/888934028824

Table I.10.1 [2/2] Snapshot of minimum achievement in reading and mathematics

		CI3[1]	Achievement at Level 2 and above in reading			Achievement at Level 2 and above in mathematics		
			Share of 15-year-old students achieving at Level 2 or above (2018)[2]	Change in this share, expressed as a percentage of the PISA target population (2009 to 2018)[3]	Change in this share, expressed as a percentage of all 15-year-olds (2009 to 2018)[3]	Share of 15-year-old students achieving at Level 2 or above (2018)[2]	Change in this share, expressed as a percentage of the PISA target population (2012 to 2018)[4]	Change in this share, expressed as a percentage of all 15-year-olds (2012 to 2018)[4]
			%	% dif.	% dif.	%	% dif.	% dif.
Partners	Albania	0.76	47.8	N.S.	N.A.	57.6	18.3	N.A.
	Argentina	0.81	47.9	N.S.	N.A.	31.0	N.S.	N.A.
	Baku (Azerbaijan)	0.46	39.6	M	M	49.3	M	M
	Belarus	0.88	76.6	M	M	70.6	M	M
	Bosnia and Herzegovina	0.82	46.3	M	M	42.4	M	M
	Brazil	0.65	50.0	N.S.	N.S.	31.9	N.S.	N.S.
	Brunei Darussalam	0.97	48.2	M	M	52.1	M	M
	B-S-J-Z (China)	0.81	94.8	M	M	97.6	M	M
	Bulgaria	0.72	52.9	N.S.	N.S.	55.6	N.S.	N.S.
	Cambodia	0.28	7.5	M	M	9.9	M	M
	Costa Rica	0.63	58.0	-9.3	N.S.	40.0	N.S.	N.S.
	Croatia	0.89	78.4	N.S.	N.A.	68.8	N.S.	N.A.
	Dominican Republic	0.73	20.9	-6.9	N.S.	9.4	N.S.	N.S.
	Ecuador	0.61	49.4	M	M	29.1	M	M
	Georgia	0.83	35.6	N.S.	N.A.	38.9	-4.0	N.A.
	Guatemala	0.47	29.9	M	M	10.6	M	M
	Honduras	0.41	29.7	M	M	15.4	M	M
	Hong Kong (China)	0.98	87.4	-4.3	N.A.	90.8	N.S.	N.A.
	Indonesia	0.85	30.1	-16.5	N.A.	28.1	N.S.	N.A.
	Jordan	0.54	58.8	N.S.	-8.6	40.7	9.2	N.S.
	Kazakhstan	0.92	35.8	N.S.	N.A.	50.9	N.S.	N.A.
	Kosovo	0.84	21.3	N.S.	N.A.	23.4	N.S.	N.A.
	Lebanon	0.87	32.2	N.S.	N.A.	40.2	N.S.	N.A.
	Macao (China)	0.88	89.2	4.1	N.A.	95.0	5.8	N.A.
	Malaysia	0.72	54.2	N.S.	N.S.	58.5	10.3	N.S.
	Malta	0.97	64.1	N.S.	N.A.	69.8	N.S.	N.A.
	Moldova	0.95	57.0	14.2	N.A.	49.7	N.S.	N.A.
	Montenegro	0.95	55.6	N.S.	N.A.	53.8	10.5	N.A.
	Morocco	0.64	26.7	M	M	24.4	M	M
	North Macedonia	0.95	44.9	15.5	N.A.	39.0	9.2	N.A.
	Panama	0.53	35.7	N.S.	N.S.	18.8	M	M
	Paraguay	M	32.2	M	M	8.3	M	M
	Peru	0.73	45.7	10.5	7.7	39.7	14.2	10.7
	Philippines	0.68	19.4	M	M	19.3	M	M
	Qatar	0.92	49.1	12.6	N.A.	46.3	15.9	N.A.
	Romania	0.73	59.2	N.S.	N.S.	53.4	N.S.	N.S.
	Russia	0.94	77.9	5.3	N.A.	78.4	N.S.	N.A.
	Saudi Arabia	0.85	47.6	M	M	27.3	M	M
	Senegal	0.29	8.7	M	M	7.7	M	M
	Serbia	0.88	62.3	N.S.	N.A.	60.3	N.S.	N.A.
	Singapore	0.95	88.8	N.S.	N.A.	92.9	N.S.	N.A.
	Chinese Taipei	0.92	82.2	N.S.	N.A.	86.0	N.S.	N.A.
	Thailand	0.72	40.5	-16.7	-12.3	47.3	N.S.	N.S.
	Ukraine	0.87	74.1	M	M	64.1	M	M
	United Arab Emirates	0.92	57.1	N.S.	N.A.	54.5	N.S.	N.A.
	Uruguay	0.77	58.1	N.S.	N.A.	49.3	N.S.	N.A.
	Zambia	0.36	5.0	M	M	2.3	M	M

1. CI3: Coverage Index 3, corresponding to the proportion of 15-year-olds who are represented by the PISA sample. For Paraguay, Coverage Index 3 is reported as missing; see Chapter 11 in the *PISA for Development Technical Report* (OECD, 2018[3]) (https://www.oecd.org/pisa/pisa-for-development/pisafordevelopment2018technicalreport/PISA_D_Chapter_11_SamplingOutcomes.pdf, accessed on 28 August 2019).
2. Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia: data refer to 2017 and were collected as part of the PISA for Development assessment.
3. Austria, OECD average-37 and United Arab Emirates: 2012 to 2018; Dominican Republic, Kosovo, Lebanon and North Macedonia: 2015 to 2018.
4. Dominican Republic, Georgia, Kosovo, Lebanon, Malta, Moldova and North Macedonia: 2015 to 2018.

N.S.: not significant.
N.A.: not applicable (Coverage Index 3 is above 0.75).
M: missing due to data availability.

Source: OECD, PISA 2018 Database, Table I.B1.49.

StatLink https://doi.org/10.1787/888934028824

Changes in technology and society will continue to influence the demand for skills and the contexts in which adults and young people will use their competence in literacy and numeracy. Nevertheless, PISA Level 2 proficiency, which is used in PISA reports (including this one) to identify low-achieving students in reading and mathematics, can represent the “minimum proficiency level” referred to in Target 4.1.[2] This definition of minimum proficiency was accepted by the TCG.

Table I.10.1 shows, for each country and economy, the proportion of 15-year-old students who attained proficiency Level 2 in reading and mathematics in 2018 and, where available, the change in this proportion since 2009 (for reading) and 2012 (for mathematics). In countries with marked changes in enrolment rates over these periods, it is important to account for these differences when measuring progress towards greater inclusion and quality in education. For this reason, an alternative measure of progress towards this target is also included for countries where the coverage rate of PISA samples was below 75% (meaning that 25% or more of 15-year-olds were either out of school, in school but enrolled below 6th grade, or excluded from PISA) in 2018. This alternative measure neutralises the impact of changes in enrolment rates (or, more precisely, in the coverage rate of the PISA sample with respect to the 15-year-old population) by computing the proportion of students who scored above the minimum proficiency level not only amongst students represented by PISA samples, but amongst the entire population of 15-year-olds.

The share of 15-year-old students, in grade 7 and above, who reached a minimum level of proficiency in reading (i.e. at least Level 2 on the PISA scale) ranged from close to 90% (in Beijing, Shanghai, Jiangsu and Zhejiang [China], Estonia, Macao [China] and Singapore) to less than 10% in Cambodia, Senegal and Zambia, countries that participated in the PISA for Development assessment in 2017 (Table I.10.1). In mathematics, the share of 15-year-old students who attained minimum levels of proficiency (Level 2 and above on the PISA scale) varied even more – between 98% in Beijing, Shanghai, Jiangsu and Zhejiang (China) and 2% in Zambia. On average across OECD countries, 77% of 15-year-olds attained the minimum level of proficiency in reading, and 76% attained that level in mathematics. These numbers show that, in 2018, all countries still have some way to go towards reaching the global goals for quality education.

Table I.10.1 also shows those countries that have made significant progress over the past decade towards the objective of ensuring that all children reach minimum levels of proficiency in reading and mathematics by the end of lower secondary education. The share of students who scored above minimum levels in reading grew by more than 10 percentage points in the Republic of Moldova, the Republic of North Macedonia, Peru and Qatar, in particular; in mathematics, similarly large increases in the share of students performing above minimum proficiency in reading were observed in Albania, Malaysia, Montenegro, Peru and Qatar.

For countries where the share of students represented by PISA (Coverage Index 3) corresponds to less than 75% of all 15-year-olds (often as a result of early dropout, late or discontinuous enrolment, and grade-retention in primary school), an alternative measure of progress towards the target is also presented in Table I.10.1. Instead of comparing shares of students over time, this alternative measure relates the number of students who performed above the minimum level of proficiency to the total population of 15-year-olds in the country. Assessed in this way, progress can result either from increases in the share of students who performed above the target or, if this share remains stable, from increases in the proportion of 15-year-olds who were in school in grade 7 or above.

This measure combines aspects related to the “quantity” of schooling (i.e. the share of 15-year-olds who are enrolled in school, in grade 7 and above) with measures of the “quality” of education outcomes (i.e. the share of students who scored above the minimum level of proficiency). In doing so, the measure encourages countries that, in 2018, still had comparatively low educational attainment amongst a significant share of young people, to work not only to improve the quality of teaching and learning in school, but also to make their secondary education systems more inclusive. By this measure, Turkey should also be counted amongst the countries that made rapid progress towards Target 4.1 over the past decade.

The children who are expected to meet the target of minimum proficiency in core subjects by 2030 have already been born. For most countries, the numbers presented in Table I.10.1 represent more than a baseline against which future progress can be measured. They represent an urgent call to action to ensure that, as these children progress through the various stages of education – from pre-primary, to primary to secondary education – there are social and education policies in place to support families, communities and schools in their efforts to help all children realise their potential.

SDG Target 4.5

Target 4.5 is dedicated to equity: “By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations”. This target is cross-cutting by nature and encompasses all types of inequality across all education outcomes.

PISA is helping countries monitor progress in reducing disparities, particularly with respect to the attainment of minimum levels of proficiency (SDG Target 4.1). The TCG on the Indicators for SDG 4 identifies “parity indices” as the main measure to be used in monitoring inequalities (see Annex A3). Amongst the many dimensions of inequality and vulnerability identified for Indicator 4.5.1, PISA can help monitor gender disparities and inequalities related to family resources, through statistics based on the PISA index of economic, social and cultural status.[3]

Equity in education is analysed in detail in *PISA 2018 Results (Volume II): Where All Students Can Succeed* (OECD, 2019[4]), which contains a wide set of indicators on within-country inequalities in learning outcomes, and on the fairness and inclusiveness of education systems. Table I.10.2 shows, for each country/economy, only a single indicator of gender and socio-economic inequalities in minimum proficiency. This indicator, called the parity index, compares the share of 15-year-old students who reached at least Level 2 performance across two groups of students that differ in some background characteristics. The parity index varies between 0 and 2. It is equal to 1 if the share of 15-year-old students scoring above minimum levels is the same for both groups (no disparity).[4] For example: if the share of girls scoring above Level 2 is 40%, and the share of boys is 50%, the gender parity index is 0.8 (40%/50%). Conversely, if the share of girls is 50% and the share of boys is 40%, the gender parity index is 1.2 (2 – 40%/50%). Values close to 1 indicate either a small percentage-point difference between the two shares or, for a given percentage-point difference, a higher average share. In other words, the parity index is sensitive both to differences in performance and to overall levels of performance.

Table I.10.2 [1/2] **Snapshot of disparities in minimum achievement in reading and mathematics**

		Gender disparities in minimum achievement (Parity index[1] for girls, compared to boys)		Socio-economic disparities in minimum achievement (Parity index[1] for disadvantaged students, compared to advantaged students[2])	
		Reading (2018)[3]	Mathematics (2018)[3]	Reading (2018)[3]	Mathematics (2018)[3]
		Parity index	Parity index	Parity index	Parity index
OECD	Australia	1.11	0.99	0.76	0.71
	Austria	1.13	0.99	0.70	0.70
	Belgium	1.08	0.97	0.68	0.67
	Canada	1.09	1.00	0.85	0.81
	Chile	1.13	0.93	0.63	0.39
	Colombia	1.07	0.75	0.44	0.34
	Czech Republic	1.13	1.01	0.68	0.66
	Denmark	1.11	1.01	0.78	0.80
	Estonia	1.07	1.00	0.90	0.88
	Finland	1.13	1.04	0.85	0.80
	France	1.11	1.00	0.70	0.64
	Germany	1.10	1.00	0.71	0.68
	Greece	1.22	1.04	0.63	0.57
	Hungary	1.12	0.98	0.58	0.55
	Iceland	1.19	1.07	0.73	0.76
	Ireland	1.07	1.00	0.84	0.78
	Israel	1.22	1.09	0.57	0.53
	Italy	1.11	0.97	0.72	0.69
	Japan	1.09	1.00	0.80	0.85
	Korea	1.08	1.01	0.82	0.80
	Latvia	1.16	1.00	0.78	0.78
	Lithuania	1.18	1.05	0.68	0.65
	Luxembourg	1.13	0.97	0.58	0.59
	Mexico	1.11	0.88	0.47	0.44
	Netherlands	1.13	1.02	0.73	0.78
	New Zealand	1.11	0.99	0.75	0.70
	Norway	1.16	1.05	0.81	0.78
	Poland	1.11	1.02	0.81	0.78
	Portugal	1.10	1.00	0.71	0.65
	Slovak Republic	1.18	1.01	0.56	0.57
	Slovenia	1.16	1.01	0.79	0.77
	Spain	m	m	0.73	0.68
	Sweden	1.11	1.02	0.77	0.73
	Switzerland	1.12	0.99	0.68	0.76
	Turkey	1.14	0.97	0.71	0.65
	United Kingdom	1.07	0.97	0.81	0.76
	United States	1.09	0.98	0.76	0.62
	OECD average	1.12	0.99	0.72	0.68

1. Values of the parity index below 1 indicate a disparity in favour of the second group (boys, or advantaged students). Values of the parity index above 1 indicate a disparity in favour of the first group (girls, or disadvantaged students). Values equal to 1 indicate equal shares amongst both groups.
2. Socio-economically advantaged students are students in the top quarter of the PISA index of economic, social and cultural status (ESCS) in their own country/economy. Socio-economically disadvantaged students are students in the bottom quarter of the PISA index of economic, social and cultural status (ESCS) in their own country/economy.
3. Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia: data refer to 2017 and were collected as part of the PISA for Development assessment.

Source: OECD, PISA 2018 Database, Table I.B1.50.

StatLink https://doi.org/10.1787/888934028843

...

Table I.10.2 [2/2] Snapshot of disparities in minimum achievement in reading and mathematics

	Gender disparities in minimum achievement (Parity index[1] for girls, compared to boys)		Socio-economic disparities in minimum achievement (Parity index[1] for disadvantaged students, compared to advantaged students[2])	
	Reading (2018)[3]	Mathematics (2018)[3]	Reading (2018)[3]	Mathematics (2018)[3]
	Parity index	Parity index	Parity index	Parity index
Partners				
Albania	1.35	1.06	0.51	0.75
Argentina	1.11	0.78	0.36	0.20
Baku (Azerbaijan)	1.27	0.94	0.57	0.63
Belarus	1.13	0.99	0.61	0.54
Bosnia and Herzegovina	1.30	1.01	0.50	0.45
Brazil	1.20	0.88	0.45	0.26
Brunei Darussalam	1.23	1.07	0.40	0.47
B-S-J-Z (China)	1.03	1.00	0.92	0.96
Bulgaria	1.27	1.03	0.40	0.45
Cambodia	1.31	0.84	0.22	0.19
Costa Rica	1.11	0.80	0.50	0.37
Croatia	1.16	0.98	0.80	0.68
Dominican Republic	1.37	0.94	0.23	0.12
Ecuador	1.09	0.71	0.41	0.27
Georgia	1.37	1.04	0.39	0.40
Guatemala	1.15	0.84	0.25	0.10
Honduras	1.11	0.66	0.35	0.20
Hong Kong (China)	1.10	1.03	0.89	0.89
Indonesia	1.31	1.13	0.39	0.37
Jordan	1.35	1.01	0.60	0.52
Kazakhstan	1.31	1.00	0.56	0.75
Kosovo	1.34	0.87	0.40	0.42
Lebanon	1.22	0.99	0.25	0.37
Macao (China)	1.06	1.00	0.96	0.96
Malaysia	1.23	1.07	0.45	0.48
Malta	1.26	1.11	0.64	0.62
Moldova	1.26	1.02	0.44	0.38
Montenegro	1.24	0.94	0.63	0.60
Morocco	1.31	0.97	0.33	0.32
North Macedonia	1.41	1.09	0.45	0.39
Panama	1.16	0.82	0.27	0.15
Paraguay	1.12	0.56	0.34	0.15
Peru	1.13	0.85	0.29	0.24
Philippines	1.34	1.11	0.11	0.16
Qatar	1.41	1.21	0.46	0.40
Romania	1.22	0.98	0.47	0.40
Russia	1.12	1.00	0.79	0.76
Saudi Arabia	1.44	1.12	0.42	0.29
Senegal	1.11	0.86	0.28	0.36
Serbia	1.22	1.01	0.62	0.60
Singapore	1.07	1.01	0.83	0.86
Chinese Taipei	1.08	1.02	0.77	0.79
Thailand	1.38	1.16	0.41	0.54
Ukraine	1.16	0.97	0.63	0.54
United Arab Emirates	1.33	1.09	0.48	0.43
Uruguay	1.17	0.93	0.46	0.39
Zambia	1.45	1.26	0.04	0.04

1. Values of the parity index below 1 indicate a disparity in favour of the second group (boys, or advantaged students). Values of the parity index above 1 indicate a disparity in favour of the first group (girls, or disadvantaged students). Values equal to 1 indicate equal shares amongst both groups.
2. Socio-economically advantaged students are students in the top quarter of the PISA index of economic, social and cultural status (ESCS) in their own country/economy. Socio-economically disadvantaged students are students in the bottom quarter of the PISA index of economic, social and cultural status (ESCS) in their own country/economy.
3. Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia: data refer to 2017 and were collected as part of the PISA for Development assessment.

Source: OECD, PISA 2018 Database, Table I.B1.50.

StatLink https://doi.org/10.1787/888934028843

Table I.10.2 shows that gender disparities in minimum proficiency are often in favour of girls in reading (as indicated by values of the parity index above 1) and of boys in mathematics. In both subjects, these disparities tend to be limited, as indicated by parity indices between 0.85 and 1.15.

In contrast, socio-economic disparities are more systematic across subjects and only a few countries/economies had limited disparities in above-minimum proficiency related to socio-economic status. These include Beijing, Shanghai, Jiangsu and Zhejiang (China), Estonia, Hong Kong (China) and Macao (China). Across OECD countries, the average parity index for socio-economic differences in performance above minimum levels (i.e. at Level 2 and above) was 0.72 in reading and 0.68 in mathematics. This means that, on average across OECD countries, there were only about seven socio-economically disadvantaged students who scored above the minimum proficiency level in reading or mathematics for every 10 advantaged students who scored above that level.[5] Disparities were even wider in several low- and middle-income countries, including Cambodia, the Dominican Republic, Guatemala, Panama, Peru, the Philippines and Zambia, where the socio-economic parity index was lower than 0.30 in both reading and mathematics.

Other thematic targets and means of implementation

PISA also provides useful data for monitoring some thematic indicators that are relevant to Target 4.7 ("ensure all learners acquire knowledge and skills needed to promote sustainable development [...]"), particularly through its assessments of science (15-year-olds' ability to engage with science-related issues, and with the ideas of science, as reflective citizens) and global competence (their ability to understand and appreciate the perspectives and world views of others). PISA indicators of students' global competence are discussed in *PISA 2018 Results (Volume VI): Are Students Ready to Thrive in Global Societies?* (OECD, forthcoming[5]).

In addition, data on the context in which students learn enable countries to monitor two of the three "means of implementation" for SDG 4. In particular, PISA data can be used to monitor the quality of education facilities (Target 4.a: "facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all"); and the supply of qualified teachers (Target 4.c: "teachers that are in sufficient number and adequately trained, qualified, motivated and supported").[6] PISA indicators related to resources, including teachers, are discussed in *PISA 2018 Results (Volume V): Effective Policies, Successful Schools* (OECD, forthcoming[6]).

HOW PISA AND THE OECD ARE HELPING COUNTRIES BUILD NATIONAL SYSTEMS FOR MONITORING LEARNING GOALS

Through participation in PISA, countries can also enhance their capacity to develop relevant data to monitor national and international learning targets at different levels of education. While most countries that have participated in PISA already have adequate systems in place, that is not true for many low- and middle-income countries. To this end, the OECD PISA for Development initiative not only aimed to expand the coverage of the international assessment to include more middle- and low-income countries, it also offered these countries assistance in building their national assessment and data-collection systems. These capacity-building components of the PISA programme are now offered to all new countries joining PISA for its 2021 or 2024 cycle.

Countries that took part in PISA for Development prepared for their participation through a process that began with an analysis of their capacity to implement PISA and make use of PISA data, and included planning to strengthen that capacity. Countries were supported by the OECD and its contractors at each stage of the assessment cycle. This process helped countries overcome two potential barriers to participation in PISA: a lack of capacity to implement the assessment and a lack of experience in using PISA data and results. To overcome the latter obstacle, the OECD and its contractors offered training and assistance in data analysis, the interpretation of PISA results, report writing and communication.

During the analysis phase of the project, analysts confirmed that the test instruments measured what they purported to measure, and that the population statistics derived from tests and questionnaires could be compared internationally and used to monitor global learning goals. National analysis teams therefore could use the assessment results for a report that included relevant comparisons to inform decisions concerning national policies. Each country's report highlighted main messages from the results as well as policy options to pursue to improve learning outcomes.[7]

Other OECD data, such as those derived from the Survey of Adult Skills (a product of the OECD Programme for the International Assessment of Adult Competencies [PIAAC]) and the OECD Teaching and Learning International Survey (TALIS), provide a solid evidence base for monitoring education goals more widely. PIAAC, in particular, is the principal source of data for measuring progress towards SDG Target 4.6 – adult literacy and numeracy. OECD data complement and inspire national data systems and promote peer learning, as countries can compare their experiences in implementing policies through their own analyses, or through reviews and reports co-ordinated by the OECD.

Notes

1. Evidence supporting the comparability of PISA for Development results with results from the PISA pen-and-pencil tests can be found in Chapter 12 of the *PISA for Development Technical Report* (OECD, 2018[3]).

2. Level 2 proficiency is already used as a normative benchmark in many countries. For example, the European Union's strategic framework for co-operation in education and training (known as ET 2020), established in 2009, states: "By 2020, the share of low-achieving 15-year-olds in reading, mathematics and science should be less than 15%" (as measured by the proportion of 15-year-old students performing below Level 2 in PISA) (European Council, 2009[7]).

3. In many countries, PISA samples are stratified according to geography and can also be used to monitor disparities related to location (e.g. region or city). These comparisons are not included here, as the location categories must be defined differently for each country.

4. If the share in the first group is smaller than that in the second group, the parity index is defined as the ratio of the share in the first group (e.g. girls) divided by the share in the second group (e.g. boys). If the share in the second group is smaller, the parity index is defined as two minus the inverse ratio.

5. It is possible to interpret socio-economic parity indices in this way because socio-economic groups (advantaged and disadvantaged) are defined to be of equal size, each comprising one quarter of a country's/economy's 15-year-old students.

6. Data from two other OECD programmes, the Teaching and Learning International Survey (TALIS) and the Indicators of Education Systems (INES) programme, can also be used to monitor progress towards Target 4.c.

7. PISA for Development capacity-needs analyses and national reports can be accessed through www.oecd.org/pisa/pisa-for-development/pisa-for-development-documentation.htm (accessed on 13 July 2019).

References

European Council (2009), *Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ("ET 2020")*, Office for Official Publications of the European Communities, pp. C 119/2-10, http://dx.doi.org/10.3000/17252423.C_2009.119.eng. [7]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed,* PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [4]

OECD (2018), *PISA for Development Technical Report*, OECD Publishing, Paris, http://www.oecd.org/pisa/pisa-for-development/pisafordevelopment2018technicalreport/ (accessed on 3 October 2019). [3]

OECD (forthcoming), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris. [6]

OECD (forthcoming), *PISA 2018 Results (Volume VI): Are Students Ready to Thrive in Global Societies?* PISA, OECD Publishing, Paris. [5]

UNESCO Institute for Statistics (2019), *Quick Guide to Education Indicators for SDG 4*, UNESCO Institute for Statistics, https://unesdoc.unesco.org/ark:/48223/pf0000265396 (accessed on 13 July 2019). [1]

United Nations Statistics Division (2019), *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*, https://unstats.un.org/sdgs/indicators/indicators-list/ (accessed on 13 July 2019). [2]

PISA

ANNEX A

PISA 2018 technical background

All figures and tables in Annex A are available on line

ANNEX A1
The construction of proficiency scales and of indices from the student context questionnaire

PROFICIENCY SCALES FOR READING, MATHEMATICS AND SCIENCE

Proficiency scores in reading, mathematics and science are based on student responses to items that represent the assessment framework for each domain (see Chapter 2). While different students saw different questions, the test design, which ensured a significant overlap of items across different forms, made it possible to construct proficiency scales that are common to all students for each domain. In general, the PISA frameworks assume that a single continuous scale can be used to report overall proficiency in a domain; but this assumption is further verified during scaling (see below).

PISA proficiency scales are constructed using item-response-theory models, in which the likelihood that the test-taker responds correctly to any question is a function of the question's characteristics (see below) and of the test-taker's position on the scale. In other words, the test-taker's proficiency is associated with a particular point on the scale that indicates the likelihood that he or she responds correctly to any question. Higher values on the scale indicate greater proficiency, which is equivalent to a greater likelihood of responding correctly to any question. A description of the modelling technique used to construct proficiency scales can be found in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

In the item-response-theory models used in PISA, the task characteristics are summarised by two parameters that represent task difficulty and task discrimination. The first parameter, task difficulty, is the point on the scale where there is at least a 50% probability of a correct response by students who score at or above that point; higher values correspond to more difficult items. For the purpose of describing proficiency levels that represent mastery, PISA often reports the difficulty of a task as the point on the scale where there is at least a 62% probability of a correct response by students who score at or above that point.[1]

The second parameter, task discrimination, represents the rate at which the proportion of correct responses increases as a function of student proficiency. For an idealised highly discriminating item, close to 0% of students respond correctly if their proficiency is below the item difficulty, and close to 100% of students respond correctly as soon as their proficiency is above the item difficulty. In contrast, for weakly discriminating items, the probability of a correct response still increases as a function of student proficiency, but only gradually.

A single continuous scale can therefore show both the difficulty of questions and the proficiency of test-takers (see Figure I.2.1). By showing the difficulty of each question on this scale, it is possible to locate the level of proficiency in the domain that the question demands. By showing the proficiency of test-takers on the same scale, it is possible to describe each test-taker's level of skill or literacy by the type of tasks that he or she can perform correctly most of the time.

Estimates of student proficiency are based on the kinds of tasks that students are expected to perform successfully. This means that students are likely to be able to successfully answer questions located at or below the level of difficulty associated with their own position on the scale. Conversely, they are unlikely to be able to successfully answer questions above the level of difficulty associated with their position on the scale.[2]

The higher a student's proficiency level is located above a given test question, the more likely is he or she to be able to answer the question successfully. The discrimination parameter for this particular test question indicates how quickly the likelihood of a correct response increases. The further the student's proficiency is located below a given question, the less likely is he or she to be able to answer the question successfully. In this case, the discrimination parameter indicates how fast this likelihood decreases as the distance between the student's proficiency and the question's difficulty increases.

How many scales per domain? Assessing the dimensionality of PISA domains

PISA frameworks for reading, mathematics and science assume that a single continuous scale can summarise performance in each domain for all countries. This assumption is incorporated in the item-response-theory model used in PISA. Violations of this assumption therefore result in model misfit, and can be assessed by inspecting fit indices.

After the field trial, initial estimates of model fit for each item, and for each country and language group, provide indications about the plausibility of the uni-dimensionality assumption and about the equivalence of scales across countries. These initial estimates are used to refine the item set used in each domain: problematic items are sometimes corrected (e.g. if a translation error is detected); and coding and scoring rules can be amended (e.g. to suppress a partial-credit score that affected coding reliability, or to combine responses to two or more items when the probability of a correct response to one question appears to depend on the correct answer to an earlier question). Items can also be deleted after the field trial. Deletions are carefully balanced so that the set of retained items continues to provide a good balance of all aspects of the framework.

After the main study, the estimates of model fit are mainly used to refine the scaling model (some limited changes to the scoring rules and item deletions can also be considered). In response to earlier criticisms (Kreiner and Christensen, 2013[2]; Oliveri and von Davier, 2013[3]) and to take advantage of the increased computational resources available, PISA, in its 2015 cycle, moved to a more flexible item-response-theory model. This model allows items to vary not only in difficulty, but in their ability to discriminate between high and low performance. It also assigns country- and language-specific characteristics to items that do not fit the model for the particular item and language (see Annex A6 and OECD, forthcoming[1]). This "tailoring" of the measurement model makes it possible to improve model fit considerably, while retaining the desired level of comparability across countries and the interpretation of scales through a single set of proficiency descriptors.

With the 2015 assessment, PISA also introduced an additional test of dimensionality to confirm that "trend" and "new" items can be reported on the same scale. Using the international dataset, this test compares fit statistics for a model assuming uni-dimensionality with fit statistics for a model that assumes that "trend" and "new" items represent two distinct continuous traits. In 2015, for science, and then again in 2018, for reading, this test confirmed that a uni-dimensional model for "trend" and "new" items fits the data almost as well as a two-dimensional model, and that a uni-dimensional scale is more reliable than separate scales for "trend" and "new" items. This evidence was interpreted as showing that a single coherent scale can represent the constructs of science and reading in 2018 (OECD, forthcoming[1]).

Despite the evidence in favour of a uni-dimensional scale, for the "major" domain (i.e. reading in PISA 2018) PISA nevertheless provides multiple estimates of performance, in addition to the overall scale, through so-called "subscales". Subscales represent different framework dimensions and provide a more nuanced picture of performance in a domain. Subscales within a domain are usually highly correlated across students (thus supporting the assumption that a coherent overall scale can be formed by combining items across subscales). Despite this high correlation, interesting differences in performance across subscales can often be observed at aggregate levels (across countries, across education systems within countries, or between boys and girls).

How reporting scales are set and linked across multiple assessments

The reporting scale for each domain was originally established when the domain was the major focus of assessment in PISA for the first time: PISA 2000 for reading, PISA 2003 for mathematics and PISA 2006 for science.

The item-response-theory models used in PISA describe the relationship between student proficiency, item difficulty and item discrimination, but do not set a measurement unit for any of these parameters. In PISA, this measurement unit is chosen the first time a reporting scale is established. The score of "500" on the scale is defined as the average proficiency of students across OECD countries; "100 score points" is defined as the standard deviation (a measure of the variability) of proficiency across OECD countries.[3]

To enable the measurement of trends, achievement data from successive assessments are reported on the same scale. It is possible to report results from different assessments on the same scale because in each assessment PISA retains a significant number of items from previous PISA assessments. These are known as trend items. All items used to assess mathematics and science in 2018, and a significant number of items used to assess reading (72 out of 244), were developed and already used in earlier assessments (see Tables I.A5.1 and I.A5.3). Their difficulty and discrimination parameters were therefore already estimated in previous PISA assessments.

The answers to the trend questions from students in earlier PISA cycles, together with the answers from students in PISA 2018, were both considered when scaling PISA 2018 data to determine student proficiency, item difficulty and item discrimination. In particular, when scaling PISA 2018 data, item parameters for new items were freely estimated, but item parameters for trend items were initially fixed to their PISA 2015 values, which, in turn, were based on a concurrent calibration involving response data from multiple cycles (OECD, 2017[4]). All constraints on trend item parameters were evaluated and, in some cases, released in order to better describe students' response patterns. See the *PISA 2018 Technical Report* (OECD, forthcoming[1]) for details.

The extent to which the item characteristics estimated during the scaling of PISA 2018 data differ from those estimated in previous calibrations is summarised in the "link error", a quantity (expressed in score points) that reflects the uncertainty in comparing PISA results over time. A link error of zero indicates a perfect match in the parameters across calibrations, while a non-zero link error indicates that the relative difficulty of certain items or the ability of certain items to discriminate between high and low achievers has changed over time, introducing greater uncertainty in trend comparisons.

INDICES FROM THE STUDENT CONTEXT QUESTIONNAIRE

In addition to scale scores representing performance in reading, mathematics and science, this volume uses indices derived from the PISA student questionnaires to contextualise PISA 2018 results or to estimate trends that account for demographic changes over time. The following indices and database variables are used:

- Student age (database variable: AGE)
- Student gender (ST004)
- Immigrant background (IMMIG)
- Language spoken at home (ST022)
- The PISA index of economic, social and cultural status (ESCS)

For a description of how these indices were constructed, see Annex A1 in *PISA 2018 Results (Volume II): Where all Students Can Succeed* (OECD, 2019[5]) and the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Chapter 1 also reports changes, over time, in time spent using the Internet (2012 and 2018), in the proportion of students having access to various digital devices, in the number of devices available at home, and in students' reading habits and attitudes towards reading (2009 and 2018).

Most of these analyses report proportions of particular answer categories in the student questionnaire or in the ICT familiarity questionnaire, which was optional for countries. In a few cases, some answer categories were combined (e.g. "agree" and "strongly agree") prior to conducting the analysis; these simple recodes are indicated in column headers and in notes under Tables I.B1.54-I.B1.59.

In addition, three indices were used for the analysis of time spent using the Internet, in Tables I.B1.51-I.B1.53. The indices of time spent using the Internet were constructed from students' answers to the following questions, which were included in the optional ICT familiarity questionnaire:

- During a typical weekday, for how long do you use the Internet at school? (IC005)
- During a typical weekday, for how long do you use the Internet outside of school? (IC006)
- On a typical weekend day, for how long do you use the Internet outside of school? (IC007)

Students were allowed to respond in intervals of: no time; between 1-30 minutes per day; between 31-60 minutes per day; between 1 hour and 2 hours per day; between 2 hours and 4 hours per day; between 4 hours and 6 hours per day; and more than 6 hours per day. To build the indices of time spent using the Internet, these responses were converted to the smallest number of minutes in the interval (0, 1, 31, 61, 121, 241 or 361, respectively). As such, the indices represent lower bounds of the time spent on the Internet reported by each student.

Notes

1. This definition of task difficulty, referred to as RP62 in the *PISA 2018 Technical Report* (OECD, forthcoming[1]), is used in particular to classify assessment items into proficiency levels (see Chapter 5). The choice of a probability of 62%, rather than of 50%, sets the bar for mastery of a particular level of proficiency significantly above chance levels, including for simple multiple-choice response formats. In the typical parametrisation of the two-parameters IRT-model used by PISA, RP62 values depend on both model parameters.

2. "Unlikely", in this context, refers to a probability below 62%.

3. The standard deviation of 100 score points corresponds to the standard deviation in a pooled sample of students from OECD countries, where each national sample is equally weighted.

References

Kreiner, S. and **K. Christensen** (2013), "Analyses of Model Fit and Robustness. A New Look at the PISA Scaling Model Underlying Ranking of Countries According to Reading Literacy", *Psychometrika*, Vol. 79/2, pp. 210-231, http://dx.doi.org/10.1007/s11336-013-9347-z. [2]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [5]

OECD (2017), *PISA 2015 Technical Report*, OECD Publishing, Paris, http://www.oecd.org/pisa/data/2015-technical-report/ (accessed on 31 July 2017). [4]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

Oliveri, M. and **M. von Davier** (2013), "Toward Increasing Fairness in Score Scale Calibrations Employed in International Large-Scale Assessments", *International Journal of Testing*, Vol. 14/1, pp. 1-21, http://dx.doi.org/10.1080/15305058.2013.825265. [3]

ANNEX A2
The PISA target population, the PISA samples and the definition of schools

Exclusions and coverage ratios

WHO IS THE PISA TARGET POPULATION?

PISA 2018 assessed the cumulative outcomes of education and learning at a point at which most young people are still enrolled in formal education – when they are 15 years old.

Any international survey of education must guarantee the comparability of its target population across nations. One way to do this is to assess students at the same grade level. However, differences between countries in the nature and extent of pre-primary education and care, the age at entry into formal schooling, and the institutional structure of education systems do not allow for a definition of internationally comparable grade levels.

Other international assessments have defined their target population by the grade level that provides maximum coverage of a particular age cohort. However, this method is particularly sensitive to the distribution of students across age and grade levels; small changes in this distribution can lead to the selection of different target grades, even within the same country over different PISA cycles. There also may be differences across countries in whether students who are older or younger than the desired age cohort are represented in the modal grade, further rendering such grade-level-based samples difficult to compare.

To overcome these problems, PISA uses an age-based definition of its target population, one that is not tied to the institutional structures of national education systems. PISA assesses students who are aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months[1] at the beginning of the assessment period, plus or minus an allowed 1-month variation, and who are enrolled in an educational institution[2] at grade 7 or higher.[3] All students who met these criteria were eligible to sit the PISA assessment in 2018, regardless of the type of educational institution in which they were enrolled and whether they were enrolled in full-time or part-time education. This also allows PISA to evaluate students shortly before they are faced with major life choices, such as whether to continue with education or enter the workforce.

Hence, PISA makes statements about the knowledge and skills of a group of individuals who were born within a comparable reference period, but who may have undergone different educational experiences both in and outside of school. These students may be distributed over different ranges of grades (both in terms of the specific grade levels and the spread in grade levels) in different countries, or over different tracks or streams. It is important to consider these differences when comparing PISA results across countries. In addition, differences in performance observed when students are 15 may disappear later on if students' experiences in education converge over time.

If a country's mean scores in reading, mathematics or science are significantly higher than those of another country, it cannot automatically be inferred that schools or particular parts of the education system in the first country are more effective than those in the second. However, one can legitimately conclude that it is the cumulative impact of learning experiences in the first country, starting in early childhood and up to the age of 15, and including all experiences, whether they be at school, home or elsewhere, that have resulted in the better outcomes of the first country in the subjects that PISA assesses.[4]

The PISA target population does not include residents of a country who attend school in another country. It does, however, include foreign nationals who attend school in the country of assessment.

To accommodate countries that requested grade-based results for the purpose of national analyses, PISA 2018 provided a sampling option to supplement age-based sampling with grade-based sampling.

HOW WERE STUDENTS CHOSEN?

The accuracy of the results from any survey depends on the quality of the information drawn from those surveyed as well as on the sampling procedures. Quality standards, procedures, instruments and verification mechanisms were developed for PISA that ensured that national samples yielded comparable data and that the results could be compared across countries with confidence. Experts from the PISA Consortium selected the samples for most participating countries/economies and monitored the sample-selection process closely in those countries that selected their own samples.

Most PISA samples were designed as two-stage stratified samples.[5] The first stage sampled schools in which 15-year-old students may be enrolled. Schools were sampled systematically with probabilities proportional to the estimated size of their (eligible) 15-year-old population. At least 150 schools[6] were selected in each country, although the requirements for national analyses often demanded a larger sample. Replacement schools for each sampled school were simultaneously identified, in case an originally sampled school chose not to participate in PISA 2018.

The second stage of the selection process sampled students within sampled schools. Once schools were selected, a list of each sampled school's 15-year-old students was prepared. From this list, 42 students were then selected with equal probability (all 15-year-old students were selected if fewer than 42 were enrolled). The target number of students who were to be sampled in a school could deviate from 42 but could not fall below 20.

Data-quality standards in PISA required minimum participation rates for schools as well as for students. These standards were established to minimise the potential for bias resulting from non-response. Indeed, it was likely that any bias resulting from non-response would be negligible – i.e. typically smaller than the sampling error – in countries that met these standards.

At least 85% of the schools initially selected to take part in the PISA assessment were required to agree to conduct the test. Where the initial response rate of schools was between 65% and 85%, however, an acceptable school-response rate could still be achieved through the use of replacement schools. Inherent in this procedure was a risk of introducing bias, if replacement schools differed from initially sampled schools along dimensions other than those considered for sampling. Participating countries and economies were therefore encouraged to persuade as many of the schools in the original sample as possible to participate.

Schools with a student participation rate of between 25% and 50% were not considered to be participating schools, but data (from both the cognitive assessment and questionnaire) from these schools were included in the database and contributed to the various estimates. Data from schools with a student participation rate of less than 25% were excluded from the database.

In PISA 2018, five countries and economies – Hong Kong (China) (69%), Latvia (82%), New Zealand (83%), the United Kingdom (73%) and the United States (65%) – did not meet the 85% threshold, but met the 65% threshold, amongst schools initially selected to take part in the PISA assessment. Upon replacement, Hong Kong (China) (79%), the United Kingdom (87%) and the United States (76%) still failed to reach an acceptable participation rate.[7] Amongst the schools initially selected before replacement, the Netherlands (61%) did not meet the 65% school response-rate threshold, but it reached a response rate of 87% upon replacement. However, these were not considered to be major issues as, for each of these countries/economies, additional non-response analyses showed that there were limited differences between schools that did participate and the full set of schools originally drawn in the sample.[8] Data from these jurisdictions were hence considered to be largely comparable with, and were therefore reported together with, data from other countries/economies.

PISA 2018 also required that at least 80% of the students chosen within participating schools participated themselves. This threshold was calculated at the national level and did not have to be met in each participating school. Follow-up sessions were required in schools where too few students had participated in the original assessment sessions. Student-participation rates were calculated over all original schools; and also over all schools, whether original or replacement schools. Students who participated in either the original or in any follow-up assessment sessions were counted in these participation rates; those who attended only the questionnaire session were included in the international database and contributed to the statistics presented in this publication if they provided at least a description of their father's or mother's occupation.

This 80% threshold was met in every country/economy except Portugal, where only 76% of students who were sampled actually participated. The high level of non-responding students could lead to biased results, e.g. if students who did not respond were more likely to be low-performing students. This was indeed the case in Portugal, but a non-response analysis based on data from a national mathematics assessment in the country showed that the upward bias of Portugal's overall results was likely small enough to preserve comparability over time and with other countries. Data from Portugal was therefore reported along with data from the countries/economies that met this 80% student-participation threshold.

Table I.A2.6 shows the response rate for students and schools, before and after replacement.

- **Column 1** shows the weighted participation rate of schools before replacement; it is equivalent to Column 2 divided by Column 3 (multiplied by 100 to give a percentage).
- **Column 2** shows the number of responding schools before school replacement, weighted by student enrolment.
- **Column 3** shows the number of sampled schools before school replacement, weighted by student enrolment. This includes both responding and non-responding schools.
- **Column 4** shows the unweighted number of responding schools before school replacement.

- **Column 5** shows the unweighted number of sampled schools before school replacement, including both responding and non-responding schools.
- **Columns 6 to 10** repeat Columns 1 to 5 for schools *after* school replacement, i.e. after non-responding schools were replaced by the replacement schools identified during the initial sampling procedure.
- **Columns 11 to 15** repeat Columns 6 to 10 but for *students* in schools after school replacement. Note that the weighted and unweighted numbers of students sampled (Columns 13 and 15) include students who were assessed and those who should have been assessed but who were absent on the day of assessment. Furthermore, as mentioned above, any students in schools where the student response rate was less than 50% were not considered to be attending participating schools, and were thus excluded from Columns 14 and 15 (and, similarly, from Columns 4, 5, 9 and 10).

WHAT PROPORTION OF 15-YEAR-OLDS DOES PISA REPRESENT?

All countries and economies attempted to maximise the coverage of 15-year-olds enrolled in education in their national samples, including students enrolled in special-education institutions.

The sampling standards used in PISA only permitted countries and economies to exclude up to a total of 5% of the relevant population (i.e. 15-year-old students enrolled in school at grade 7 or higher) either by excluding schools or excluding students within schools. All but 16 countries and economies – Sweden (11.09%), Israel (10.21%), Luxembourg (7.92%), Norway (7.88%), Canada (6.87%), New Zealand (6.78%), Switzerland (6.68%), the Netherlands (6.24%), Cyprus (5.99%), Iceland (5.99%), Kazakhstan (5.87%), Australia (5.72%), Denmark (5.70%), Turkey (5.66%), the United Kingdom (5.45%) and Estonia (5.03%) – achieved this standard, and in 28 countries and economies, the overall exclusion rate was less than 2% (Table I.A2.1) When language exclusions[9] were accounted for (i.e. removed from the overall exclusion rate), Estonia and Iceland no longer had exclusion rates greater than 5%. More details can be found in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Exclusions that should remain within the above limits include both:

- at the school level:
  - schools that were geographically inaccessible or where the administration of the PISA assessment was not considered feasible
  - schools that provided teaching only for students in the categories defined under "within-school exclusions", such as schools for the blind.

  The percentage of 15-year-olds enrolled in such schools had to be less than 2.5% of the nationally desired target population (0.5% maximum for the former group and 2% maximum for the latter group). The magnitude, nature and justification of school-level exclusions are documented in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).
- at the student level:
  - students with an intellectual disability, i.e. a mental or emotional disability resulting in the student being so cognitively delayed that he/she could not perform in the PISA testing environment
  - students with a functional disability, i.e. a moderate to severe permanent physical disability resulting in the student being unable to perform in the PISA testing environment
  - students with limited assessment-language proficiency. These students were unable to read or speak any of the languages of assessment in the country at a sufficient level and unable to overcome such a language barrier in the PISA testing environment, and were typically students who had received less than one year of instruction in the language of assessment
  - other exclusions, a category defined by the PISA national centres in individual participating countries and approved by the PISA international consortium
  - students taught in a language of instruction for the major domain for which no materials were available.

  Students could not be excluded solely because of low proficiency or common disciplinary problems. The percentage of 15-year-olds excluded within schools had to be less than 2.5% of the national desired target population.

Although exceeding the exclusion rate limit of 5% (Table I.A2.1), data from the 16 countries and economies listed above were all deemed to be acceptable for the reasons listed below. In particular, all of these reasons were accepted by a data-adjudication panel to allow for the reliable comparison of PISA results across countries and economies and across time; thus the data from these countries were reported together with data from other countries/economies.

- In Australia, Canada, Denmark, Luxembourg, New Zealand and Norway, exclusion rates remained close to those observed in previous cycles. In the United Kingdom, exclusion rates were also above 5% but have decreased markedly across cycles.
- In Cyprus, Iceland, Kazakhstan, the Netherlands and Switzerland, exclusions increased but remained close to the 5% limit. The increase could be largely attributed to a marked increase in students who were excluded within schools due to intellectual or functional disabilities. Moreover, in the Netherlands, some 17% of students were not excluded but assigned to UH (*une heure*) booklets, which were intended for students with special education needs. As these booklets did not cover the domain of financial literacy (see *PISA 2018 Results [Volume V]: Are Students Smart about Money?*, OECD, forthcoming[2]), the effective exclusion rate for the Netherlands in financial literacy was over 20%. This resulted in a strong upward bias in the country mean and other population statistics in that domain. Data from the Netherlands in financial literacy are not comparable with data from other education systems; but data from the Netherlands in the core PISA subjects were still deemed to be largely comparable.
- The higher exclusion rate in Turkey was likely the result of a higher school-level exclusion rate due to a particular type of non-formal educational institution that was not listed (and hence not excluded) in 2015 but was listed and excluded in 2018.
- The higher exclusion rate in Israel was the result of a higher school-level exclusion rate due to the lack of participation by a particular type of boys' school. These schools were considered to be non-responding schools in cycles up to 2015 but were treated as school-level exclusions in 2018.
- Sweden had the highest exclusion rate: 11.07%. It is believed that this increase in the exclusion rate was due to a large and temporary increase in immigrant and refugee inflows, although because of Swedish data-collection laws, this could not be explicitly stated in student-tracking forms. Instead, students confronted with language barriers were classified as being excluded "for other reasons", as were students with intellectual and functional disabilities. It is expected that the exclusion rate will decrease to previous levels in future cycles of PISA, as such inflows stabilise or shrink.[10]

Table I.A2.1 describes the target population of the countries participating in PISA 2018. Further information on the target population and the implementation of PISA sampling standards can be found in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

- **Column 1** shows the total number of 15-year-olds according to the most recent available information, which in most countries and economies means from 2017, the year before the assessment.
- **Column 2** shows the number of 15-year-olds enrolled in school in grade 7 or above, which is referred to as the "eligible population".
- **Column 3** shows the national desired target population. Countries and economies were allowed to exclude up to 0.5% of students *a priori* from the eligible population, essentially for practical reasons. The following *a priori* exclusions exceed this limit but were agreed with the PISA Consortium:
  - Canada excluded 1.17% of its population: students living in the Yukon, Northwest Territories and Nunavut, and Aboriginal students living on reserves
  - Chile excluded 0.05% of its population: students living on Easter Island, the Juan Fernandez Archipelago and Antarctica
  - Cyprus excluded 0.10% of its population: students attending schools on the northern part of the island
  - the Philippines excluded 2.42% of its population: students living in the Autonomous Region in Muslim Mindanao
  - Saudi Arabia excluded 7.59% of its population: students living in the regions of Najran and Jizan
  - Ukraine excluded 0.37% of its population: some students attending schools in the Donetsk and Luhansk regions
  - the United Arab Emirates excluded 0.04% of its population: home-schooled students.
- **Column 4** shows the number of students enrolled in schools that were excluded from the national desired target population, either from the sampling frame or later in the field during data collection. In other words, these are school-level exclusions.
- **Column 5** shows the size of the national desired target population after subtracting the students enrolled in excluded schools. This column is obtained by subtracting Column 4 from Column 3.
- **Column 6** shows the percentage of students enrolled in excluded schools. This is obtained by dividing Column 4 by Column 3 and multiplying by 100.
- **Column 7** shows the number of students who participated in PISA 2018. Note that in some cases, this number does not account for 15-year-olds assessed as part of additional national options.

- **Column 8** shows the weighted number of participating students, i.e. the number of students in the nationally defined target population that the PISA sample represents.
- **Column 9** shows the total number of students excluded within schools. In each sampled school, all eligible students – namely, those 15 years of age, regardless of grade – were listed, and a reason for the exclusion was provided for each student who was to be excluded from the sample. These reasons are further described and classified into specific categories in Table I.A2.4.
- **Column 10** shows the weighted number of students excluded within schools, i.e. the overall number of students in the national defined target population represented by the number of students from the sample excluded within schools. This weighted number is also described and classified by exclusion categories in Table I.A2.4.
- **Column 11** shows the percentage of students excluded within schools. This is equivalent to the weighted number of excluded students (Column 10) divided by the weighted number of excluded and participating students (the sum of Columns 8 and 10), multiplied by 100.
- **Column 12** shows the overall exclusion rate, which represents the weighted percentage of the national desired target population excluded from PISA either through school-level exclusions or through the exclusion of students within schools. It is equivalent to the school-level exclusion rate (Column 6) plus the product of the within-school exclusion rate and 1 minus the school-level exclusion rate expressed as a decimal (Column 6 divided by 100).[11]
- **Column 13** shows an index of the extent to which the national desired target population was covered by the PISA sample. As mentioned above, 16 countries/economies fell below the coverage of 95%. This is also known as Coverage Index 1.
- **Column 14** shows an index of the extent to which 15-year-olds *enrolled in school* were covered by the PISA sample. The index, also known as Coverage Index 2, measures the overall proportion of the national enrolled population that is covered by the non-excluded portion of the student sample, and takes into account both school- and student-level exclusions. Values close to 100 indicate that the PISA sample represents the entire (grade 7 and higher) education system as defined for PISA 2018. This is calculated in a similar manner to Column 13; however, the total enrolled population of 15-year-olds in grade 7 or above (Column 2) is used as a base instead of the national desired target population (Column 3).
- **Column 15** shows an index of the coverage of the 15-year-old population. The index is the weighted number of participating students (Column 8) divided by the total population of 15-year-old students (Column 1). This is also known as Coverage Index 3.

A high level of coverage contributes to the comparability of the assessment results. For example, even assuming that the excluded students would have systematically scored worse than those who participated, and that this relationship is moderately strong, an exclusion rate on the order of 5% would likely lead to an overestimation of national mean scores of less than 5 score points on the PISA scale (where the standard deviation is 100 score points).[12]

DEFINITION OF SCHOOLS

In some countries, subunits within schools were sampled instead of schools, which may affect the estimate of the between-school variance. In Austria, the Czech Republic, Germany, Hungary, Japan, Romania and Slovenia, schools with more than one programme of study were split into the units delivering these programmes. In the Netherlands, locations were listed as sampling units. In the Flemish Community of Belgium, each campus (or implantation) of a multi-campus school was sampled independently, whereas the larger administrative unit of a multi-campus school was sampled as a whole in the French Community of Belgium.

In Argentina, Australia, Colombia and Croatia, each campus of a multi-campus school was sampled independently. Schools in the Basque Country of Spain that were divided into sections by language of instruction were split into these linguistic sections for sampling. International schools in Luxembourg were split into two sampling units: one for students who were instructed in a language for which testing material was available,[13] and one for students who were instructed in a language for which no testing material was available (and who were hence excluded).

Some schools in the United Arab Emirates were sampled as a whole unit, while others were split by curriculum and sometimes by gender. Due to reorganisation, some schools in Sweden were split into two parts, each part with its own principal. Some schools in Portugal were organised into clusters where all units in a cluster shared the same teachers and principal; each of these clusters constituted a single sampling unit.

THE DISTRIBUTION OF PISA STUDENTS ACROSS GRADES

Students assessed in PISA 2018 were enrolled in various grade levels. The percentage of students at each grade level is presented, by country, in Table I.A2.8 and Table I.A2.9, and by gender within each country in Table I.A2.12 and Table I.A2.13.

Table I.A2.1 [1/4] **PISA target populations and samples**

		Population and sample information						
		Total population of 15-year-olds	Total enrolled population of 15-year-olds at grade 7 or above	Total in national desired target population	Total school-level exclusions	Total in national desired target population after all school exclusions and before within-school exclusions	School-level exclusion rate (%)	Number of participating students
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
OECD	Australia	288 195	284 687	284 687	5 610	279 077	1.97	14 273
	Austria	84 473	80 108	80 108	603	79 505	0.75	6 802
	Belgium	126 031	122 808	122 808	1 877	120 931	1.53	8 475
	Canada	388 205	400 139	395 448	7 950	387 498	2.01	22 653
	Chile	239 492	215 580	215 470	2 151	213 319	1.00	7 621
	Colombia	856 081	645 339	645 339	950	644 389	0.15	7 522
	Czech Republic	92 013	90 835	90 835	1 510	89 325	1.66	7 019
	Denmark	68 313	67 414	67 414	653	66 761	0.97	7 657
	Estonia	12 257	12 120	12 120	413	11 707	3.41	5 316
	Finland	58 325	57 552	57 552	496	57 056	0.86	5 649
	France	828 196	798 480	798 480	13 732	784 748	1.72	6 308
	Germany	739 792	739 792	739 792	15 448	724 344	2.09	5 451
	Greece	102 868	100 203	100 203	1 266	98 937	1.26	6 403
	Hungary	96 838	91 297	91 297	1 992	89 305	2.18	5 132
	Iceland	4 232	4 177	4 177	35	4 142	0.84	3 294
	Ireland	61 999	61 188	61 188	59	61 129	0.10	5 577
	Israel	136 848	128 419	128 419	10 613	117 806	8.26	6 623
	Italy	616 185	544 279	544 279	748	543 531	0.14	11 785
	Japan	1 186 849	1 159 226	1 159 226	27 743	1 131 483	2.39	6 109
	Korea	517 040	517 040	517 040	2 489	514 551	0.48	6 650
	Latvia	17 977	17 677	17 677	692	16 985	3.92	5 303
	Lithuania	27 075	25 998	25 998	494	25 504	1.90	6 885
	Luxembourg	6 291	5 952	5 952	156	5 796	2.62	5 230
	Mexico	2 231 751	1 697 100	1 697 100	8 013	1 689 087	0.47	7 299
	Netherlands	208 704	204 753	204 753	10 347	194 406	5.05	4 765
	New Zealand	59 700	58 131	58 131	857	57 274	1.47	6 173
	Norway	60 968	60 794	60 794	852	59 942	1.40	5 813
	Poland	354 020	331 850	331 850	6 853	324 997	2.07	5 625
	Portugal	112 977	110 732	110 732	709	110 023	0.64	5 932
	Slovak Republic	51 526	50 100	50 100	587	49 513	1.17	5 965
	Slovenia	17 501	18 236	18 236	337	17 899	1.85	6 401
	Spain	454 168	436 560	436 560	2 368	434 192	0.54	35 943
	Sweden	108 622	107 824	107 824	1 492	106 332	1.38	5 504
	Switzerland	80 590	78 059	78 059	3 227	74 832	4.13	5 822
	Turkey	1 218 693	1 038 993	1 038 993	43 928	995 065	4.23	6 890
	United Kingdom	703 991	697 603	697 603	1 315	64 076	2.01	13 818
	United States	4 133 719	4 058 637	4 058 637	24 757	4 033 880	0.61	4 838

Notes: For a full explanation of the details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.1 [2/4] **PISA target populations and samples**

		Population and sample information						
		Total population of 15-year-olds	Total enrolled population of 15-year-olds at grade 7 or above	Total in national desired target population	Total school-level exclusions	Total in national desired target population after all school exclusions and before within-school exclusions	School-level exclusion rate (%)	Number of participating students
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Partners	Albania	36 955	30 160	30 160	0	30 160	0.00	6 359
	Argentina	702 788	678 151	678 151	5 597	672 554	0.83	11 975
	Baku (Azerbaijan)	43 798	22 672	22 672	454	22 218	2.00	6 827
	Belarus	89 440	82 580	82 580	1 440	81 140	1.74	5 803
	Bosnia and Herzegovina	35 056	32 313	32 313	243	32 070	0.75	6 480
	Brazil	3 132 463	2 980 084	2 980 084	74 772	2 905 312	2.51	10 691
	Brunei Darussalam	7 081	7 384	7 384	0	7 384	0.00	6 828
	B-S-J-Z (China)	1 221 746	1 097 296	1 097 296	33 279	1 064 017	3.03	12 058
	Bulgaria	66 499	51 674	51 674	388	51 286	0.75	5 294
	Costa Rica	72 444	58 789	58 789	0	58 789	0.00	7 221
	Croatia	39 812	30 534	30 534	409	30 125	1.34	6 609
	Cyprus	8 285	8 285	8 277	138	8 139	1.67	5 503
	Dominican Republic	192 198	148 033	148 033	2 755	145 278	1.86	5 674
	Georgia	46 605	41 750	41 750	1 018	40 732	2.44	5 572
	Hong Kong (China)	51 935	51 328	51 328	643	50 685	1.25	6 037
	Indonesia	4 439 086	3 684 980	3 684 980	3 892	3 681 088	0.11	12 098
	Jordan	212 777	132 291	132 291	90	132 201	0.07	8 963
	Kazakhstan	230 646	230 018	230 018	9 814	220 204	4.27	19 507
	Kosovo	30 494	27 288	27 288	87	27 201	0.32	5 058
	Lebanon	61 979	59 687	59 687	1 300	58 387	2.18	5 614
	Macao (China)	4 300	3 845	3 845	14	3 831	0.36	3 775
	Malaysia	537 800	455 358	455 358	3 503	451 855	0.77	6 111
	Malta	4 039	4 056	4 056	37	4 019	0.91	3 363
	Moldova	29 716	29 467	29 467	78	29 389	0.26	5 367
	Montenegro	7 484	7 432	7 432	40	7 392	0.54	6 666
	Morocco	601 250	415 806	415 806	8 292	407 514	1.99	6 814
	North Macedonia	18 812	18 812	18 812	298	18 514	1.59	5 569
	Panama	72 084	60 057	60 057	585	59 472	0.97	6 270
	Peru	580 690	484 352	484 352	10 483	473 869	2.16	6 086
	Philippines	2 063 564	1 734 997	1 692 950	42 290	1 650 660	2.50	7 233
	Qatar	16 492	16 408	16 408	245	16 163	1.49	13 828
	Romania	203 940	171 685	171 685	4 653	167 032	2.71	5 075
	Russia	1 343 738	1 339 706	1 339 706	48 114	1 291 592	3.59	7 608
	Saudi Arabia	418 788	406 768	375 914	8 940	366 974	2.38	6 136
	Serbia	69 972	66 729	66 729	1 175	65 554	1.76	6 609
	Singapore	46 229	45 178	45 178	552	44 626	1.22	6 676
	Chinese Taipei	246 260	240 241	240 241	1 978	238 263	0.82	7 243
	Thailand	795 130	696 833	696 833	10 014	686 819	1.44	8 633
	Ukraine	351 424	321 833	320 636	8 352	312 284	2.60	5 998
	United Arab Emirates	59 275	59 203	59 178	847	58 331	1.43	19 277
	Uruguay	50 965	46 768	46 768	0	46 768	0.00	5 263
	Viet Nam	1 332 000	1 251 842	1 251 842	6 169	1 245 673	0.49	5 377

Notes: For a full explanation of the details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.1 [3/4] **PISA target populations and samples**

		Population and sample information					Coverage indices		
		Weighted number of participating students	Number of excluded students	Weighted number of excluded students	Within-school exclusion rate (%)	Overall exclusion rate (%)	Coverage Index 1: Coverage of national desired population	Coverage Index 2: Coverage of national enrolled population	Coverage Index 3: Coverage of 15-year-old population
		(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD	**Australia**	257 779	716	10 249	3.82	5.72	0.943	0.943	0.894
	Austria	75 077	117	1 379	1.80	2.54	0.975	0.975	0.889
	Belgium	118 025	45	494	0.42	1.94	0.981	0.981	0.936
	Canada	335 197	1 481	17 496	4.96	6.87	0.931	0.920	0.863
	Chile	213 832	68	2 029	0.94	1.93	0.981	0.980	0.893
	Colombia	529 976	28	1 812	0.34	0.49	0.995	0.995	0.619
	Czech Republic	87 808	1	11	0.01	1.67	0.983	0.983	0.954
	Denmark	59 967	444	3 009	4.78	5.70	0.943	0.943	0.878
	Estonia	11 414	96	195	1.68	5.03	0.950	0.950	0.931
	Finland	56 172	157	1 491	2.59	3.42	0.966	0.966	0.963
	France	756 477	56	6 644	0.87	2.58	0.974	0.974	0.913
	Germany	734 915	42	4 847	0.66	2.73	0.973	0.973	0.993
	Greece	95 370	52	798	0.83	2.08	0.979	0.979	0.927
	Hungary	86 754	75	1 353	1.54	3.68	0.963	0.963	0.896
	Iceland	3 875	209	212	5.19	5.99	0.940	0.940	0.916
	Ireland	59 639	257	2 370	3.82	3.91	0.961	0.961	0.962
	Israel	110 645	152	2 399	2.12	10.21	0.898	0.898	0.809
	Italy	521 223	93	3 219	0.61	0.75	0.992	0.992	0.846
	Japan	1 078 921	0	0	0.00	2.39	0.976	0.976	0.909
	Korea	455 544	7	378	0.08	0.56	0.994	0.994	0.881
	Latvia	15 932	23	62	0.38	4.29	0.957	0.957	0.886
	Lithuania	24 453	95	360	1.45	3.32	0.967	0.967	0.903
	Luxembourg	5 478	315	315	5.44	7.92	0.921	0.921	0.871
	Mexico	1 480 904	44	11 457	0.77	1.24	0.988	0.988	0.664
	Netherlands	190 281	78	2 407	1.25	6.24	0.938	0.938	0.912
	New Zealand	53 000	443	3 016	5.38	6.78	0.932	0.932	0.888
	Norway	55 566	452	3 906	6.57	7.88	0.921	0.921	0.911
	Poland	318 724	116	5 635	1.74	3.77	0.962	0.962	0.900
	Portugal	98 628	158	1 749	1.74	2.37	0.976	0.976	0.873
	Slovak Republic	44 418	12	72	0.16	1.33	0.987	0.987	0.862
	Slovenia	17 138	124	298	1.71	3.52	0.965	0.965	0.979
	Spain	416 703	747	8 951	2.10	2.63	0.974	0.974	0.918
	Sweden	93 129	681	10 163	9.84	11.09	0.889	0.889	0.857
	Switzerland	71 683	152	1 955	2.66	6.68	0.933	0.933	0.889
	Turkey	884 971	95	13 463	1.50	5.66	0.943	0.943	0.726
	United Kingdom	597 240	688	20 562	3.33	5.45	0.945	0.945	0.848
	United States	3 559 045	194	119 057	3.24	3.83	0.962	0.962	0.861

Notes: For a full explanation of the details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.1 [4/4] **PISA target populations and samples**

		Population and sample information				Coverage indices		
		Weighted number of participating students	Number of excluded students	Weighted number of excluded students	Within-school exclusion rate (%)	Overall exclusion rate (%)	Coverage Index 1: Coverage of national desired population	Coverage Index 2: Coverage of national enrolled population	Coverage Index 3: Coverage of 15-year-old population
		(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Partners	Albania	27 963	0	0	0.00	0.00	1.000	1.000	0.757
	Argentina	566 486	118	4 083	0.72	1.54	0.985	0.985	0.806
	Baku (Azerbaijan)	20 271	0	0	0.00	2.00	0.980	0.980	0.463
	Belarus	78 333	31	462	0.59	2.32	0.977	0.977	0.876
	Bosnia and Herzegovina	28 843	24	106	0.36	1.11	0.989	0.989	0.823
	Brazil	2 036 861	41	8 180	0.40	2.90	0.971	0.971	0.650
	Brunei Darussalam	6 899	53	53	0.76	0.76	0.992	0.992	0.974
	B-S-J-Z (China)	992 302	34	1 452	0.15	3.17	0.968	0.968	0.812
	Bulgaria	47 851	80	685	1.41	2.15	0.978	0.978	0.720
	Costa Rica	45 475	39	249	0.54	0.54	0.995	0.995	0.628
	Croatia	35 462	135	637	1.76	3.08	0.969	0.969	0.891
	Cyprus	7 639	201	351	4.40	5.99	0.940	0.939	0.922
	Dominican Republic	140 330	0	0	0.00	1.86	0.981	0.981	0.730
	Georgia	38 489	26	180	0.46	2.89	0.971	0.971	0.826
	Hong Kong (China)	51 101	0	0	0.00	1.25	0.987	0.987	0.984
	Indonesia	3 768 508	0	0	0.00	0.11	0.999	0.999	0.849
	Jordan	114 901	44	550	0.48	0.54	0.995	0.995	0.540
	Kazakhstan	212 229	300	3 624	1.68	5.87	0.941	0.941	0.920
	Kosovo	25 739	26	132	0.51	0.83	0.992	0.992	0.844
	Lebanon	53 726	1	8	0.02	2.19	0.978	0.978	0.867
	Macao (China)	3 799	0	0	0.00	0.36	0.996	0.996	0.883
	Malaysia	388 638	37	2 419	0.62	1.38	0.986	0.986	0.723
	Malta	3 925	56	56	1.41	2.31	0.977	0.977	0.972
	Moldova	28 252	35	207	0.73	0.99	0.990	0.990	0.951
	Montenegro	7 087	4	12	0.18	0.71	0.993	0.993	0.947
	Morocco	386 408	4	220	0.06	2.05	0.980	0.980	0.643
	North Macedonia	17 820	18	85	0.48	2.05	0.979	0.979	0.947
	Panama	38 540	24	106	0.27	1.24	0.988	0.988	0.535
	Peru	424 586	20	1 360	0.32	2.48	0.975	0.975	0.731
	Philippines	1 400 584	10	2 039	0.15	2.64	0.974	0.950	0.679
	Qatar	15 228	192	192	1.25	2.72	0.973	0.973	0.923
	Romania	148 098	24	930	0.62	3.32	0.967	0.967	0.726
	Russia	1 257 388	96	14 905	1.17	4.72	0.953	0.953	0.936
	Saudi Arabia	354 013	1	53	0.01	2.39	0.976	0.902	0.845
	Serbia	61 895	42	409	0.66	2.41	0.976	0.976	0.885
	Singapore	44 058	35	232	0.52	1.74	0.983	0.983	0.953
	Chinese Taipei	226 698	38	1 297	0.57	1.39	0.986	0.986	0.921
	Thailand	575 713	17	1 002	0.17	1.61	0.984	0.984	0.724
	Ukraine	304 855	34	1 704	0.56	3.15	0.969	0.965	0.867
	United Arab Emirates	54 403	166	331	0.60	2.03	0.980	0.979	0.918
	Uruguay	39 746	25	164	0.41	0.41	0.996	0.996	0.780
	Viet Nam	926 260	0	0	0.00	0.49	0.995	0.995	0.695

Notes: For a full explanation of the details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).
The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.2 [1/4] **Change in the enrolment of 15-year-olds in grade 7 and above (PISA 2003 through PISA 2018)**

	PISA 2018				PISA 2015				PISA 2012			
OECD	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population
Australia	288 195	284 687	257 779	0.89	282 888	282 547	256 329	0.91	291 967	288 159	250 779	0.86
Austria	84 473	80 108	75 077	0.89	88 013	82 683	73 379	0.83	93 537	89 073	82 242	0.88
Belgium	126 031	122 808	118 025	0.94	123 630	121 954	114 902	0.93	123 469	121 493	117 912	0.95
Canada	388 205	400 139	335 197	0.86	396 966	381 660	331 546	0.84	417 873	409 453	348 070	0.83
Chile	239 492	215 580	213 832	0.89	255 440	245 947	203 782	0.80	274 803	252 733	229 199	0.83
Colombia	856 081	645 339	529 976	0.62	760 919	674 079	567 848	0.75	889 729	620 422	560 805	0.63
Czech Republic	92 013	90 835	87 808	0.95	90 391	90 076	84 519	0.94	96 946	93 214	82 101	0.85
Denmark	68 313	67 414	59 967	0.88	68 174	67 466	60 655	0.89	72 310	70 854	65 642	0.91
Estonia	12 257	12 120	11 414	0.93	11 676	11 491	10 834	0.93	12 649	12 438	11 634	0.92
Finland	58 325	57 552	56 172	0.96	58 526	58 955	56 934	0.97	62 523	62 195	60 047	0.96
France	828 196	798 480	756 477	0.91	807 867	778 679	734 944	0.91	792 983	755 447	701 399	0.88
Germany	739 792	739 792	734 915	0.99	774 149	774 149	743 969	0.96	798 136	798 136	756 907	0.95
Greece	102 868	100 203	95 370	0.93	105 530	105 253	96 157	0.91	110 521	105 096	96 640	0.87
Hungary	96 838	91 297	86 754	0.90	94 515	90 065	84 644	0.90	111 761	108 816	91 179	0.82
Iceland	4 232	4 177	3 875	0.92	4 250	4 195	3 966	0.93	4 505	4 491	4 169	0.93
Ireland	61 999	61 188	59 639	0.96	61 234	59 811	59 082	0.96	59 296	57 979	54 010	0.91
Israel	136 848	128 419	110 645	0.81	124 852	118 997	117 031	0.94	118 953	113 278	107 745	0.91
Italy	616 185	544 279	521 223	0.85	616 761	567 268	495 093	0.80	605 490	566 973	521 288	0.86
Japan	1 186 849	1 159 226	1 078 921	0.91	1 201 615	1 175 907	1 138 349	0.95	1 241 786	1 214 756	1 128 179	0.91
Korea	517 040	517 040	455 544	0.88	620 687	619 950	569 106	0.92	687 104	672 101	603 632	0.88
Latvia	17 977	17 677	15 932	0.89	17 255	16 955	15 320	0.89	18 789	18 389	16 054	0.85
Lithuania	27 075	25 998	24 453	0.90	33 163	32 097	29 915	0.90	38 524	35 567	33 042	0.86
Luxembourg	6 291	5 952	5 478	0.87	6 327	6 053	5 540	0.88	6 187	6 082	5 523	0.85
Mexico	2 231 751	1 697 100	1 480 904	0.66	2 257 399	1 401 247	1 392 995	0.62	2 114 745	1 472 875	1 326 025	0.63
Netherlands	208 704	204 753	190 281	0.91	203 234	200 976	191 817	0.94	194 000	193 190	196 262	1.01
New Zealand	59 700	58 131	53 000	0.89	60 162	57 448	54 274	0.90	60 940	59 118	53 414	0.88
Norway	60 968	60 794	55 566	0.91	63 642	63 491	58 083	0.91	64 917	64 777	59 432	0.92
Poland	354 020	331 850	318 724	0.90	380 366	361 600	345 709	0.91	425 597	410 700	379 275	0.89
Portugal	112 977	110 732	98 628	0.87	110 939	101 107	97 214	0.88	108 728	127 537	96 034	0.88
Slovak Republic	51 526	50 100	44 418	0.86	55 674	55 203	49 654	0.89	59 723	59 367	54 486	0.91
Slovenia	17 501	18 236	17 138	0.98	18 078	17 689	16 773	0.93	19 471	18 935	18 303	0.94
Spain	454 168	436 560	416 703	0.92	440 084	414 276	399 935	0.91	423 444	404 374	374 266	0.88
Sweden	108 622	107 824	93 129	0.86	97 749	97 210	91 491	0.94	102 087	102 027	94 988	0.93
Switzerland	80 590	78 059	71 683	0.89	85 495	83 655	82 223	0.96	87 200	85 239	79 679	0.91
Turkey	1 218 693	1 038 993	884 971	0.73	1 324 089	1 100 074	925 366	0.70	1 266 638	965 736	866 681	0.68
United Kingdom	703 991	697 603	597 240	0.85	747 593	746 328	627 703	0.84	738 066	745 581	688 236	0.93
United States	4 133 719	4 058 637	3 559 045	0.86	4 220 325	3 992 053	3 524 497	0.84	3 985 714	4 074 457	3 536 153	0.89

Notes: Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

For Albania, Brazil, Chile, Jordan, the Netherlands, Romania, Uruguay and Viet Nam, estimates of the total population of 15-year-olds across years have been updated to align data sources with those used in 2018. Therefore, the estimates reported in this table do not match those that appear in previous PISA reports.

For Mexico, in 2015, the total population of 15-year-olds enrolled in grade 7 or above is an estimate of the target population size of the sample frame from which the 15-year-old students were selected for the PISA test. At the time Mexico provided the information to PISA, the official figure for this population was 1 573 952.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.2 [2/4] **Change in the enrolment of 15-year-olds in grade 7 and above (PISA 2003 through PISA 2018)**

		PISA 2018				PISA 2015				PISA 2012			
		Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population
Partners	**Albania**	36 955	30 160	27 963	0.76	45 667	45 163	40 896	0.90	55 099	50 157	42 466	0.77
	Argentina	702 788	678 151	566 486	0.81	718 635	578 308	394 917	0.55	684 879	637 603	545 942	0.80
	Baku (Azerbaijan)	43 798	22 672	20 271	0.46	m	m	m	m	m	m	m	m
	Belarus	89 440	82 580	78 333	0.88	m	m	m	m	m	m	m	m
	Bosnia and Herzegovina	35 056	32 313	28 843	0.82	m	m	m	m	m	m	m	m
	Brazil	3 132 463	2 980 084	2 036 861	0.65	3 379 467	2 853 388	2 425 961	0.72	3 520 371	2 786 064	2 470 804	0.70
	Brunei Darussalam	7 081	7 384	6 899	0.97	m	m	m	m	m	m	m	m
	B-S-J-Z (China)	1 221 746	1 097 296	992 302	0.81	m	m	m	m	m	m	m	m
	Bulgaria	66 499	51 674	47 851	0.72	66 601	59 397	53 685	0.81	70 188	59 684	54 255	0.77
	Costa Rica	72 444	58 789	45 475	0.63	81 773	66 524	51 897	0.63	81 489	64 326	40 384	0.50
	Croatia	39 812	30 534	35 462	0.89	45 031	35 920	40 899	0.91	48 155	46 550	45 502	0.94
	Cyprus	8 285	8 285	7 639	0.92	9 255	9 255	8 785	0.95	9 956	9 956	9 650	0.97
	Dominican Republic	192 198	148 033	140 330	0.73	193 153	139 555	132 300	0.68	m	m	m	m
	Georgia	46 605	41 750	38 489	0.83	48 695	43 197	38 334	0.79	m	m	m	m
	Hong Kong (China)	51 935	51 328	51 101	0.98	65 100	61 630	57 662	0.89	84 200	77 864	70 636	0.84
	Indonesia	4 439 086	3 684 980	3 768 508	0.85	4 534 216	3 182 816	3 092 773	0.68	4 174 217	3 599 844	2 645 155	0.63
	Jordan	212 777	132 291	114 901	0.54	196 734	121 729	108 669	0.55	153 293	125 333	111 098	0.72
	Kazakhstan	230 646	230 018	212 229	0.92	211 407	209 555	192 909	0.91	258 716	247 048	208 411	0.81
	Kosovo	30 494	27 288	25 739	0.84	31 546	28 229	22 333	0.71	m	m	m	m
	Lebanon	61 979	59 687	53 726	0.87	64 044	62 281	42 331	0.66	m	m	m	m
	Macao (China)	4 300	3 845	3 799	0.88	5 100	4 417	4 507	0.88	6 600	5 416	5 366	0.81
	Malaysia	537 800	455 358	388 638	0.72	540 000	448 838	412 524	0.76	544 302	457 999	432 080	0.79
	Malta	4 039	4 056	3 925	0.97	4 397	4 406	4 296	0.98	m	m	m	m
	Moldova	29 716	29 467	28 252	0.95	31 576	30 601	29 341	0.93	m	m	m	m
	Montenegro	7 484	7 432	7 087	0.95	7 524	7 506	6 777	0.90	8 600	8 600	7 714	0.90
	Morocco	601 250	415 806	386 408	0.64	m	m	m	m	m	m	m	m
	North Macedonia	18 812	18 812	17 820	0.95	16 719	16 717	15 847	0.95	m	m	m	m
	Panama	72 084	60 057	38 540	0.53	m	m	m	m	m	m	m	m
	Peru	580 690	484 352	424 586	0.73	580 371	478 229	431 738	0.74	584 294	508 969	419 945	0.72
	Philippines	2 063 564	1 734 997	1 400 584	0.68	m	m	m	m	m	m	m	m
	Qatar	16 492	16 408	15 228	0.92	13 871	13 850	12 951	0.93	11 667	11 532	11 003	0.94
	Romania	203 940	171 685	148 098	0.73	218 846	176 334	164 216	0.75	212 694	146 243	140 915	0.66
	Russia	1 343 738	1 339 706	1 257 388	0.94	1 176 473	1 172 943	1 120 932	0.95	1 272 632	1 268 814	1 172 539	0.92
	Saudi Arabia	418 788	406 768	354 013	0.85	m	m	m	m	m	m	m	m
	Serbia	69 972	66 729	61 895	0.88	m	m	m	m	85 121	75 870	67 934	0.80
	Singapore	46 229	45 178	44 058	0.95	48 218	47 050	46 224	0.96	53 637	52 163	51 088	0.95
	Chinese Taipei	246 260	240 241	226 698	0.92	m	m	m	m	m	m	m	m
	Thailand	795 130	696 833	575 713	0.72	895 513	756 917	634 795	0.71	982 080	784 897	703 012	0.72
	Ukraine	351 424	321 833	304 855	0.87	m	m	m	m	m	m	m	m
	United Arab Emirates	59 275	59 203	54 403	0.92	51 687	51 518	46 950	0.91	48 824	48 446	40 612	0.83
	Uruguay	50 965	46 768	39 746	0.78	53 533	43 865	38 287	0.72	54 638	46 442	39 771	0.73
	Viet Nam	1 332 000	1 251 842	926 260	0.70	1 340 000	1 032 599	874 859	0.65	1 393 000	1 091 462	956 517	0.69

Notes: Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

For Albania, Brazil, Chile, Jordan, the Netherlands, Romania, Uruguay and Viet Nam, estimates of the total population of 15-year-olds across years have been updated to align data sources with those used in 2018. Therefore, the estimates reported in this table do not match those that appear in previous PISA reports.

For Mexico, in 2015, the total population of 15-year-olds enrolled in grade 7 or above is an estimate of the target population size of the sample frame from which the 15-year-old students were selected for the PISA test. At the time Mexico provided the information to PISA, the official figure for this population was 1 573 952.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.2 [3/4] **Change in the enrolment of 15-year-olds in grade 7 and above (PISA 2003 through PISA 2018)**

		PISA 2009				PISA 2006				PISA 2003			
		Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population
OECD	Australia	286 334	269 669	240 851	0.84	270 115	256 754	234 940	0.87	268 164	250 635	235 591	0.88
	Austria	99 818	94 192	87 326	0.87	97 337	92 149	89 925	0.92	94 515	89 049	85 931	0.91
	Belgium	126 377	126 335	119 140	0.94	124 943	124 557	123 161	0.99	120 802	118 185	111 831	0.93
	Canada	430 791	426 590	360 286	0.84	426 967	428 876	370 879	0.87	398 865	399 265	330 436	0.83
	Chile	290 056	265 542	247 270	0.85	297 085	255 459	233 526	0.79	m	m	m	m
	Colombia	893 057	582 640	522 388	0.58	897 477	543 630	537 262	0.60	m	m	m	m
	Czech Republic	122 027	116 153	113 951	0.93	127 748	124 764	128 827	1.01	130 679	126 348	121 183	0.93
	Denmark	70 522	68 897	60 855	0.86	66 989	65 984	57 013	0.85	59 156	58 188	51 741	0.87
	Estonia	14 248	14 106	12 978	0.91	19 871	19 623	18 662	0.94	m	m	m	m
	Finland	66 198	66 198	61 463	0.93	66 232	66 232	61 387	0.93	61 107	61 107	57 883	0.95
	France	749 808	732 825	677 620	0.90	809 375	809 375	739 428	0.91	809 053	808 276	734 579	0.91
	Germany	852 044	852 044	766 993	0.90	951 535	1 062 920	903 512	0.95	951 800	916 869	884 358	0.93
	Greece	102 229	105 664	93 088	0.91	107 505	110 663	96 412	0.90	111 286	108 314	105 131	0.94
	Hungary	121 155	118 387	105 611	0.87	124 444	120 061	106 010	0.85	129 138	123 762	107 044	0.83
	Iceland	4 738	4 738	4 410	0.93	4 820	4 777	4 624	0.96	4 168	4 112	3 928	0.94
	Ireland	56 635	55 464	52 794	0.93	58 667	57 648	55 114	0.94	61 535	58 997	54 850	0.89
	Israel	122 701	112 254	103 184	0.84	122 626	109 370	93 347	0.76	m	m	m	m
	Italy	586 904	573 542	506 733	0.86	578 131	639 971	520 055	0.90	561 304	574 611	481 521	0.86
	Japan	1 211 642	1 189 263	1 113 403	0.92	1 246 207	1 222 171	1 113 701	0.89	1 365 471	1 328 498	1 240 054	0.91
	Korea	717 164	700 226	630 030	0.88	660 812	627 868	576 669	0.87	606 722	606 370	533 504	0.88
	Latvia	28 749	28 149	23 362	0.81	34 277	33 659	29 232	0.85	37 544	37 138	33 643	0.90
	Lithuania	51 822	43 967	40 530	0.78	53 931	51 808	50 329	0.93	m	m	m	m
	Luxembourg	5 864	5 623	5 124	0.87	4 595	4 595	4 733	1.03	4 204	4 204	4 080	0.97
	Mexico	2 151 771	1 425 397	1 305 461	0.61	2 200 916	1 383 364	1 190 420	0.54	2 192 452	1 273 163	1 071 650	0.49
	Netherlands	199 000	198 334	183 546	0.92	197 046	193 769	189 576	0.96	194 216	194 216	184 943	0.95
	New Zealand	63 460	60 083	55 129	0.87	63 800	59 341	53 398	0.84	55 440	53 293	48 638	0.88
	Norway	63 352	62 948	57 367	0.91	61 708	61 449	59 884	0.97	56 060	55 648	52 816	0.94
	Poland	482 500	473 700	448 866	0.93	549 000	546 000	515 993	0.94	589 506	569 294	534 900	0.91
	Portugal	115 669	107 583	96 820	0.84	115 426	100 816	90 079	0.78	109 149	99 216	96 857	0.89
	Slovak Republic	72 826	72 454	69 274	0.95	79 989	78 427	76 201	0.95	84 242	81 945	77 067	0.91
	Slovenia	20 314	19 571	18 773	0.92	23 431	23 018	20 595	0.88	m	m	m	m
	Spain	433 224	425 336	387 054	0.89	439 415	436 885	381 686	0.87	454 064	418 005	344 372	0.76
	Sweden	121 486	121 216	113 054	0.93	129 734	127 036	126 393	0.97	109 482	112 258	107 104	0.98
	Switzerland	90 623	89 423	80 839	0.89	87 766	86 108	89 651	1.02	83 247	81 020	86 491	1.04
	Turkey	1 336 842	859 172	757 298	0.57	1 423 514	800 968	665 477	0.47	1 351 492	725 030	481 279	0.36
	United Kingdom	786 626	786 825	683 380	0.87	779 076	767 248	732 004	0.94	768 180	736 785	698 579	0.91
	United States	4 103 738	4 210 475	3 373 264	0.82	4 192 939	4 192 939	3 578 040	0.85	3 979 116	3 979 116	3 147 089	0.79

Notes: Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

For Albania, Brazil, Chile, Jordan, the Netherlands, Romania, Uruguay and Viet Nam, estimates of the total population of 15-year-olds across years have been updated to align data sources with those used in 2018. Therefore, the estimates reported in this table do not match those that appear in previous PISA reports.

For Mexico, in 2015, the total population of 15-year-olds enrolled in grade 7 or above is an estimate of the target population size of the sample frame from which the 15-year-old students were selected for the PISA test. At the time Mexico provided the information to PISA, the official figure for this population was 1 573 952.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.2 [4/4] **Change in the enrolment of 15-year-olds in grade 7 and above (PISA 2003 through PISA 2018)**

		PISA 2009				PISA 2006				PISA 2003			
		Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population	Total population of 15-year-olds	Total population of 15-year-olds enrolled in grade 7 or above	Weighted number of participating students	Coverage Index 3: Coverage of the national 15-year-old population
Partners	Albania	55 587	42 767	34 134	0.61	m	m	m	m	m	m	m	m
	Argentina	688 434	636 713	472 106	0.69	662 686	579 222	523 048	0.79	m	m	m	m
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m
	Belarus	m	m	m	m	m	m	m	m	m	m	m	m
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m
	Brazil	3 434 101	2 654 489	2 080 159	0.61	3 439 795	2 374 044	1 875 461	0.55	3 560 650	2 359 854	1 952 253	0.55
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m
	Bulgaria	80 226	70 688	57 833	0.72	89 751	88 071	74 326	0.83	m	m	m	m
	Costa Rica	80 523	63 603	42 954	0.53	m	m	m	m	m	m	m	m
	Croatia	48 491	46 256	43 065	0.89	54 500	51 318	46 523	0.85	m	m	m	m
	Cyprus	m	m	m	m	m	m	m	m	m	m	m	m
	Dominican Republic	m	m	m	m	m	m	m	m	m	m	m	m
	Georgia	56 070	51 351	42 641	0.76	m	m	m	m	m	m	m	m
	Hong Kong (China)	85 000	78 224	75 548	0.89	77 398	75 542	75 145	0.97	75 000	72 631	72 484	0.97
	Indonesia	4 267 801	3 158 173	2 259 118	0.53	4 238 600	3 119 393	2 248 313	0.53	4 281 895	3 113 548	1 971 476	0.46
	Jordan	133 953	107 254	104 056	0.78	122 354	126 708	90 267	0.74	m	m	m	m
	Kazakhstan	281 659	263 206	250 657	0.89	m	m	m	m	m	m	m	m
	Kosovo	m	m	m	m	m	m	m	m	m	m	m	m
	Lebanon	m	m	m	m	m	m	m	m	m	m	m	m
	Macao (China)	7 500	5 969	5 978	0.80	m	m	m	m	8 318	6 939	6 546	0.79
	Malaysia	539 295	492 758	421 448	0.78	m	m	m	m	m	m	m	m
	Malta	5 152	4 930	4 807	0.93	m	m	m	m	m	m	m	m
	Moldova	47 873	44 069	43 195	0.90	m	m	m	m	m	m	m	m
	Montenegro	8 500	8 493	7 728	0.91	9 190	8 973	7 734	0.84	m	m	m	m
	Morocco	m	m	m	m	m	m	m	m	m	m	m	m
	North Macedonia	m	m	m	m	m	m	m	m	m	m	m	m
	Panama	57 919	43 623	30 510	0.53	m	m	m	m	m	m	m	m
	Peru	585 567	491 514	427 607	0.73	m	m	m	m	m	m	m	m
	Philippines	m	m	m	m	m	m	m	m	m	m	m	m
	Qatar	10 974	10 665	9 806	0.89	8 053	7 865	7 271	0.90	m	m	m	m
	Romania	220 264	152 084	151 130	0.69	312 483	241 890	223 887	0.72	m	m	m	m
	Russia	1 673 085	1 667 460	1 290 047	0.77	2 243 924	2 077 231	1 810 856	0.81	2 496 216	2 366 285	2 153 373	0.86
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m
	Serbia	85 121	75 128	70 796	0.83	88 584	80 692	73 907	0.83	m	m	m	m
	Singapore	54 982	54 212	51 874	0.94	m	m	m	m	m	m	m	m
	Chinese Taipei	m	m	m	m	m	m	m	m	m	m	m	m
	Thailand	949 891	763 679	691 916	0.73	895 924	727 860	644 125	0.72	927 070	778 267	637 076	0.69
	Ukraine	m	m	m	m	m	m	m	m	m	m	m	m
	United Arab Emirates	41 564	40 447	38 707	0.93	m	m	m	m	m	m	m	m
	Uruguay	53 801	43 281	33 971	0.63	52 119	40 815	36 011	0.69	53 948	40 023	33 775	0.63
	Viet Nam	m	m	m	m	m	m	m	m	m	m	m	m

Notes: Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

For Albania, Brazil, Chile, Jordan, the Netherlands, Romania, Uruguay and Viet Nam, estimates of the total population of 15-year-olds across years have been updated to align data sources with those used in 2018. Therefore, the estimates reported in this table do not match those that appear in previous PISA reports.

For Mexico, in 2015, the total population of 15-year-olds enrolled in grade 7 or above is an estimate of the target population size of the sample frame from which the 15-year-old students were selected for the PISA test. At the time Mexico provided the information to PISA, the official figure for this population was 1 573 952.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.4 [1/2] **Exclusions**

	Student exclusions (unweighted)						Student exclusions (weighted)					
	Number of excluded students with functional disability (Code 1)	Number of excluded students with intellectual disability (Code 2)	Number of excluded students because of language (Code 3)	Number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total number of excluded students	Number of excluded students with functional disability (Code 1)	Number of excluded students with intellectual disability (Code 2)	Number of excluded students because of language (Code 3)	Number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total number of excluded students
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
OECD Australia	69	555	92	0	0	716	1 054	7 895	1 300	0	0	10 249
Austria	7	49	61	0	0	117	77	531	771	0	0	1 379
Belgium	8	19	18	0	0	45	87	211	196	0	0	494
Canada	125	1 040	316	0	0	1 481	1 611	11 744	4 141	0	0	17 496
Chile	6	58	4	0	0	68	173	1 727	129	0	0	2 029
Colombia	4	24	0	0	0	28	346	1 466	0	0	0	1 812
Czech Republic	1	0	0	0	0	1	11	0	0	0	0	11
Denmark	15	179	88	162	0	444	98	1 453	427	1 032	0	3 009
Estonia	3	85	8	0	0	96	8	174	13	0	0	195
Finland	6	100	22	17	12	157	55	966	204	155	111	1 491
France	8	28	20	0	0	56	776	3 397	2 471	0	0	6 644
Germany	2	18	22	0	0	42	199	1 859	2 789	0	0	4 847
Greece	2	39	11	0	0	52	29	590	179	0	0	798
Hungary	5	20	4	46	0	75	77	432	67	777	0	1 353
Iceland	5	133	61	10	0	209	5	135	62	10	0	212
Ireland	39	90	45	83	0	257	367	831	420	752	0	2 370
Israel	25	87	40	0	0	152	406	1 382	611	0	0	2 399
Italy	0	0	0	93	0	93	0	0	0	3 219	0	3 219
Japan	0	0	0	0	0	0	0	0	0	0	0	0
Korea	5	1	1	0	0	7	302	74	2	0	0	378
Latvia	2	20	1	0	0	23	5	54	2	0	0	62
Lithuania	4	91	0	0	0	95	16	344	0	0	0	360
Luxembourg	5	233	77	0	0	315	5	233	77	0	0	315
Mexico	13	28	3	0	0	44	2 609	7 301	1 547	0	0	11 457
Netherlands	7	58	9	4	0	78	236	1 813	224	134	0	2 407
New Zealand	42	279	119	0	3	443	278	1 905	812	0	21	3 016
Norway	17	327	108	0	0	452	147	2 814	944	0	0	3 906
Poland	21	87	8	0	0	116	964	4 190	481	0	0	5 635
Portugal	10	139	9	0	0	158	126	1 551	73	0	0	1 749
Slovak Republic	1	8	0	3	0	12	5	50	0	18	0	72
Slovenia	13	36	75	0	0	124	20	85	193	0	0	298
Spain	39	481	227	0	0	747	423	5 400	3 128	0	0	8 951
Sweden	0	0	0	681	0	681	0	0	0	10 163	0	10 163
Switzerland	8	71	73	0	0	152	86	813	1 056	0	0	1 955
Turkey	10	46	39	0	0	95	1 248	6 389	5 825	0	0	13 463
United Kingdom	75	573	40	0	0	688	2 448	16 592	1 522	0	0	20 562
United States	38	106	39	11	0	194	25 164	62 555	24 972	6 367	0	119 057

Note: For a full explanation of other details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Exclusion codes:

Code 1: Functional disability – student has a moderate to severe permanent physical disability.

Code 2: Intellectual disability – student has a mental or emotional disability and has either been tested as cognitively delayed or is considered in the professional opinion of qualified staff to be cognitively delayed.

Code 3: Limited assessment language proficiency – student is not a native speaker of any of the languages of the assessment in the country and has been resident in the country for less than one year.

Code 4: Other reasons defined by the national centres and approved by the international centre.

Code 5: No materials available in the language of instruction.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.4 [2/2] **Exclusions**

		Student exclusions (unweighted)						Student exclusions (weighted)					
		Number of excluded students with functional disability (Code 1)	Number of excluded students with intellectual disability (Code 2)	Number of excluded students because of language (Code 3)	Number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total number of excluded students	Number of excluded students with functional disability (Code 1)	Number of excluded students with intellectual disability (Code 2)	Number of excluded students because of language (Code 3)	Number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total number of excluded students
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Partners	Albania	0	0	0	0	0	0	0	0	0	0	0	0
	Argentina	21	96	1	0	0	118	871	3 199	13	0	0	4 083
	Baku (Azerbaijan)	0	0	0	0	0	0	0	0	0	0	0	0
	Belarus	30	1	0	0	0	31	449	13	0	0	0	462
	Bosnia and Herzegovina	8	16	0	0	0	24	29	77	0	0	0	106
	Brazil	4	36	1	0	0	41	693	7 100	386	0	0	8 180
	Brunei Darussalam	9	44	0	0	0	53	9	44	0	0	0	53
	B-S-J-Z (China)	2	24	8	0	0	34	49	1 194	209	0	0	1 452
	Bulgaria	4	76	0	0	0	80	31	653	0	0	0	685
	Costa Rica	22	12	5	0	0	39	139	78	31	0	0	249
	Croatia	7	84	4	0	40	135	33	397	24	0	182	637
	Cyprus	17	143	41	0	0	201	25	250	77	0	0	351
	Dominican Republic	0	0	0	0	0	0	0	0	0	0	0	0
	Georgia	6	20	0	0	0	26	46	134	0	0	0	180
	Hong Kong (China)	0	0	0	0	0	0	0	0	0	0	0	0
	Indonesia	0	0	0	0	0	0	0	0	0	0	0	0
	Jordan	25	17	2	0	0	44	322	204	23	0	0	550
	Kazakhstan	132	157	11	0	0	300	1 673	1 617	334	0	0	3 624
	Kosovo	0	14	0	0	12	26	0	53	0	0	79	132
	Lebanon	0	1	0	0	0	1	0	8	0	0	0	8
	Macao (China)	0	0	0	0	0	0	0	0	0	0	0	0
	Malaysia	15	22	0	0	0	37	968	1 451	0	0	0	2 419
	Malta	6	48	2	0	0	56	6	48	2	0	0	56
	Moldova	4	29	2	0	0	35	25	164	18	0	0	207
	Montenegro	0	4	0	0	0	4	0	12	0	0	0	12
	Morocco	4	0	0	0	0	4	220	0	0	0	0	220
	North Macedonia	2	3	0	0	13	18	4	8	0	0	73	85
	Panama	5	18	1	0	0	24	12	91	3	0	0	106
	Peru	11	9	0	0	0	20	756	603	0	0	0	1 360
	Philippines	2	8	0	0	0	10	376	1 663	0	0	0	2 039
	Qatar	30	150	12	0	0	192	30	150	12	0	0	192
	Romania	2	19	3	0	0	24	58	700	172	0	0	930
	Russia	14	81	1	0	0	96	2 126	12 620	159	0	0	14 905
	Saudi Arabia	0	1	0	0	0	1	0	53	0	0	0	53
	Serbia	8	11	2	0	21	42	71	148	16	0	174	409
	Singapore	4	22	9	0	0	35	25	145	62	0	0	232
	Chinese Taipei	9	28	1	0	0	38	320	957	20	0	0	1 297
	Thailand	1	16	0	0	0	17	75	927	0	0	0	1 002
	Ukraine	28	6	0	0	0	34	1 389	315	0	0	0	1 704
	United Arab Emirates	16	124	26	0	0	166	26	256	49	0	0	331
	Uruguay	4	20	1	0	0	25	29	131	5	0	0	164
	Viet Nam	0	0	0	0	0	0	0	0	0	0	0	0

Note: For a full explanation of other details in this table please refer to the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Exclusion codes:

Code 1: Functional disability – student has a moderate to severe permanent physical disability.

Code 2: Intellectual disability – student has a mental or emotional disability and has either been tested as cognitively delayed or is considered in the professional opinion of qualified staff to be cognitively delayed.

Code 3: Limited assessment language proficiency – student is not a native speaker of any of the languages of the assessment in the country and has been resident in the country for less than one year.

Code 4: Other reasons defined by the national centres and approved by the international centre.

Code 5: No materials available in the language of instruction.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.6 [1/2] **Response rates**

		Initial sample – before school replacement					Final sample – after school replacement					Final sample – students within schools after school replacement				
		Weighted school participation rate before replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)	Number of responding schools (unweighted)	Number of responding and non-responding schools (unweighted)	Weighted school participation rate before replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)	Number of responding schools (unweighted)	Number of responding and non-responding schools (unweighted)	Weighted student participation rate before replacement (%)	Number of students assessed (weighted)	Number of students sampled (assessed and absent) (weighted)	Number of students assessed (unweighted)	Number of students sampled (assessed and absent) (unweighted)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD	Australia	95	264 304	278 765	734	779	96	267 078	278 765	740	779	85	210 665	247 433	14 081	16 756
	Austria	100	78 872	78 946	291	293	100	78 872	78 946	291	293	93	69 426	75 019	6 802	7 555
	Belgium	87	103 631	119 744	256	308	95	113 259	119 719	285	308	91	101 504	111 421	8 431	9 271
	Canada	86	328 935	383 699	782	914	89	339 896	383 738	804	914	84	251 025	298 737	22 440	26 252
	Chile	90	190 060	210 669	224	258	100	209 953	210 666	255	258	93	197 940	212 625	7 601	8 156
	Colombia	95	596 406	629 729	238	250	97	610 211	629 088	244	250	93	475 820	512 614	7 480	8 036
	Czech Republic	99	86 650	87 689	330	334	99	86 650	87 689	330	334	92	79 903	86 943	6 996	7 628
	Denmark	88	52 392	59 459	328	371	93	55 170	59 109	344	371	86	48 473	56 078	7 607	8 891
	Estonia	100	11 684	11 684	231	231	100	11 684	11 684	231	231	92	10 532	11 436	5 316	5 786
	Finland	99	57 420	57 710	213	214	100	57 710	57 710	214	214	93	52 102	56 124	5 649	6 084
	France	98	769 117	784 728	244	252	100	783 049	784 728	250	252	93	698 721	754 842	6 295	6 817
	Germany	96	739 666	773 082	215	226	98	759 094	773 040	221	226	90	652 025	721 258	5 431	6 036
	Greece	85	83 158	97 793	212	256	96	94 540	98 005	240	256	96	88 019	91 991	6 371	6 664
	Hungary	98	89 754	91 208	235	245	99	90 303	91 208	236	245	94	80 693	85 878	5 129	5 458
	Iceland	98	4 178	4 282	140	160	98	4 178	4 282	140	160	87	3 285	3 791	3 285	3 791
	Ireland	100	63 179	63 179	157	157	100	63 179	63 179	157	157	86	51 575	59 639	5 577	6 445
	Israel	95	109 810	115 015	164	174	100	114 896	115 108	173	174	91	99 978	110 459	6 614	7 306
	Italy	93	505 813	541 477	510	550	98	529 552	541 672	531	550	86	437 219	506 762	11 679	13 540
	Japan	89	995 577	1 114 316	175	196	93	1 041 540	1 114 316	183	196	96	971 454	1 008 286	6 109	6 338
	Korea	100	514 768	514 768	188	188	100	514 768	514 768	188	188	97	443 719	455 544	6 650	6 810
	Latvia	82	14 020	17 049	274	349	89	15 219	17 021	308	349	89	12 752	14 282	5 303	5 923
	Lithuania	100	25 370	25 467	363	364	100	25 370	25 467	363	364	93	22 614	24 405	6 885	7 421
	Luxembourg	100	5 796	5 796	44	44	100	5 796	5 796	44	44	95	5 230	5 478	5 230	5 478
	Mexico	89	1 494 409	1 670 484	268	302	96	1 599 670	1 670 484	286	302	96	1 357 446	1 412 604	7 299	7 612
	Netherlands	61	118 705	194 486	106	175	87	169 033	194 397	150	175	83	138 134	165 739	4 668	5 617
	New Zealand	83	47 335	57 316	170	208	91	52 085	57 292	189	208	83	39 801	48 214	6 128	7 450
	Norway	98	58 521	59 889	247	254	99	59 128	59 889	250	254	91	50 009	54 862	5 802	6 368
	Poland	92	302 200	329 827	222	253	99	325 266	329 756	239	253	86	267 756	311 300	5 603	6 540
	Portugal	85	92 797	108 948	233	280	91	99 760	109 168	255	280	76	68 659	90 208	5 690	7 431
	Slovak Republic	92	45 799	49 713	348	388	96	48 391	50 361	373	388	93	39 730	42 628	5 947	6 406
	Slovenia	99	17 702	17 900	337	350	99	17 744	17 900	340	350	91	15 409	16 994	6 374	7 021
	Spain	99	427 230	432 969	1 079	1 102	99	427 899	432 969	1 082	1 102	90	368 767	410 820	35 849	39 772
	Sweden	99	101 591	102 873	218	227	99	102 075	102 873	219	227	86	79 604	92 069	5 487	6 356
	Switzerland	86	68 579	79 671	201	231	99	78 808	79 213	228	231	94	67 261	71 290	5 822	6 157
	Turkey	97	947 428	975 317	181	186	100	975 317	975 317	186	186	99	873 992	884 971	6 890	6 980
	United Kingdom	73	496 742	681 510	399	538	87	590 558	682 212	461	538	83	427 944	514 975	13 668	16 443
	United States	65	2 516 631	3 874 298	136	215	76	2 960 088	3 873 842	162	215	85	2 301 006	2 713 513	4 811	5 686

StatLink https://doi.org/10.1787/888934028862

Table I.A2.6 [2/2] **Response rates**

	Initial sample – before school replacement					Final sample – after school replacement					Final sample – students within schools after school replacement				
	Weighted school participation rate before replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)	Number of responding schools (unweighted)	Number of responding and non-responding schools (unweighted)	Weighted school participation rate before replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)	Number of responding schools (unweighted)	Number of responding and non-responding schools (unweighted)	Weighted student participation rate before replacement (%)	Number of students assessed (weighted)	Number of students sampled (assessed and absent) (weighted)	Number of students assessed (unweighted)	Number of students sampled (assessed and absent) (unweighted)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Partners															
Albania	97	29 234	30 163	322	336	97	29 260	30 163	323	336	98	26 611	27 081	6 333	6 438
Argentina	95	626 740	658 143	439	458	96	629 651	658 143	445	458	86	467 613	541 981	11 836	13 532
Baku (Azerbaijan)	93	18 730	20 040	181	197	100	20 249	20 249	197	197	89	18 049	20 312	6 827	7 607
Belarus	100	79 623	79 623	234	234	100	79 623	79 623	234	234	97	76 321	78 333	5 803	5 963
Bosnia and Herzegovina	100	31 025	31 058	212	213	100	31 051	31 051	213	213	96	27 562	28 843	6 480	6 781
Brazil	87	2 483 766	2 862 749	547	638	93	2 649 165	2 858 009	586	638	89	1 683 080	1 894 398	10 606	11 956
Brunei Darussalam	100	6 681	6 681	55	55	100	6 681	6 681	55	55	99	6 828	6 899	6 828	6 899
B-S-J-Z (China)	96	1 030 427	1 068 463	355	362	99	1 062 001	1 068 486	361	362	99	978 803	986 556	12 058	12 156
Bulgaria	96	48 095	50 164	191	199	99	49 568	50 145	197	199	93	44 003	47 275	5 294	5 673
Costa Rica	100	58 843	58 843	205	205	100	58 843	58 843	205	205	97	44 179	45 522	7 221	7 433
Croatia	97	28 382	29 188	178	183	100	29 177	29 177	183	183	92	32 632	35 462	6 609	7 190
Cyprus	98	7 946	8 122	90	99	98	7 946	8 122	90	99	93	6 975	7 472	5 503	5 890
Dominican Republic	96	138 500	143 842	225	235	100	143 816	143 816	235	235	90	126 090	140 330	5 674	6 328
Georgia	99	40 450	40 814	321	326	99	40 542	40 810	322	326	95	36 366	38 226	5 572	5 874
Hong Kong (China)	69	34 976	50 371	120	174	79	39 765	50 608	136	174	85	34 219	40 108	5 706	6 692
Indonesia	99	3 623 573	3 647 226	398	399	99	3 623 573	3 647 226	398	399	96	3 570 441	3 733 024	12 098	12 570
Jordan	100	123 056	123 056	313	313	100	123 056	123 056	313	313	98	112 213	114 901	8 963	9 172
Kazakhstan	100	220 344	220 344	616	616	100	220 344	220 344	616	616	99	210 226	212 229	19 507	19 721
Kosovo	94	25 768	27 304	203	224	97	26 324	27 269	211	224	96	23 902	24 845	5 058	5 259
Lebanon	94	54 392	58 119	302	320	98	56 652	58 093	313	320	91	47 855	52 453	5 614	6 154
Macao (China)	100	3 830	3 830	45	45	100	3 830	3 830	45	45	99	3 775	3 799	3 775	3 799
Malaysia	99	445 667	450 371	189	191	100	450 371	450 371	191	191	97	378 791	388 638	6 111	6 264
Malta	100	3 997	3 999	50	51	100	3 997	3 999	50	51	86	3 363	3 923	3 363	3 923
Moldova	100	29 054	29 054	236	236	100	29 054	29 054	236	236	98	27 700	28 252	5 367	5 474
Montenegro	99	7 242	7 299	60	61	100	7 280	7 280	61	61	96	6 822	7 087	6 666	6 912
Morocco	99	404 138	406 348	178	179	100	406 348	406 348	179	179	97	375 677	386 408	6 814	7 011
North Macedonia	100	18 489	18 502	117	120	100	18 489	18 502	117	120	92	16 467	17 808	5 569	5 999
Panama	94	54 475	57 873	241	260	97	56 455	58 002	251	260	90	34 060	37 944	6 256	7 058
Peru	99	455 964	460 276	336	342	100	460 276	460 276	342	342	99	419 329	425 036	6 086	6 170
Philippines	99	1 551 977	1 560 748	186	187	100	1 560 748	1 560 748	187	187	97	1 359 350	1 400 584	7 233	7 457
Qatar	100	16 163	16 163	188	188	100	16 163	16 163	188	188	91	13 828	15 228	13 828	15 228
Romania	98	157 747	160 607	167	170	100	160 607	160 607	170	170	98	144 688	148 098	5 075	5 184
Russia	100	1 354 843	1 355 318	264	265	100	1 354 843	1 355 318	264	265	96	1 209 339	1 257 352	7 608	7 911
Saudi Arabia	99	362 426	364 675	233	235	100	364 291	364 620	234	235	97	343 747	353 702	6 136	6 320
Serbia	97	62 037	63 877	183	190	99	63 448	63 877	187	190	94	57 342	61 233	6 609	7 062
Singapore	97	43 138	44 691	161	167	98	43 738	44 569	164	167	95	40 960	43 290	6 646	7 019
Chinese Taipei	97	232 563	238 821	186	193	99	236 227	239 027	189	193	95	211 796	223 812	7 196	7 584
Thailand	100	691 460	691 460	290	290	100	691 460	691 460	290	290	99	568 456	575 713	8 633	8 739
Ukraine	98	301 552	308 245	244	250	100	308 163	308 163	250	250	96	291 850	304 855	5 998	6 263
United Arab Emirates	99	57 891	58 234	754	760	99	57 891	58 234	754	760	96	51 517	53 904	19 265	20 191
Uruguay	97	44 528	46 032	183	189	99	45 745	46 018	188	189	87	34 333	39 459	5 247	6 026
Viet Nam	100	1 116 404	1 116 404	151	151	100	1 116 404	1 116 404	151	151	99	914 874	926 260	5 377	5 445

StatLink https://doi.org/10.1787/888934028862

Table I.A2.8 [1/2] **Percentage of students at each grade level**

		All students													
		7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above		Information unavailable	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	Australia	0.0	c	0.1	(0.0)	11.5	(0.4)	81.0	(0.5)	7.4	(0.4)	0.0	(0.0)	0.0	c
	Austria	0.4	(0.1)	6.8	(0.4)	44.5	(0.7)	48.1	(0.8)	0.2	(0.1)	0.0	c	0.0	c
	Belgium	0.3	(0.1)	6.1	(0.4)	26.7	(0.7)	63.3	(0.8)	1.3	(0.1)	0.0	c	2.3	(0.3)
	Canada	0.3	(0.1)	1.0	(0.2)	9.7	(0.3)	87.7	(0.3)	1.1	(0.1)	0.1	(0.0)	0.0	c
	Chile	1.0	(0.2)	4.4	(0.5)	20.6	(0.7)	68.5	(0.9)	5.6	(0.3)	0.0	c	0.0	c
	Colombia	4.4	(0.4)	11.3	(0.5)	22.8	(0.6)	43.0	(0.8)	18.5	(0.7)	0.0	c	0.0	c
	Czech Republic	0.6	(0.2)	3.3	(0.4)	48.5	(1.2)	47.5	(1.3)	0.0	c	0.0	c	0.0	c
	Denmark	0.1	(0.0)	16.3	(0.5)	81.7	(0.5)	1.7	(0.3)	0.0	c	0.1	(0.1)	0.0	c
	Estonia	0.4	(0.1)	21.8	(0.6)	76.4	(0.6)	1.3	(0.2)	0.0	(0.0)	0.0	c	0.0	c
	Finland	0.3	(0.1)	13.9	(0.4)	85.6	(0.5)	0.2	(0.1)	0.0	c	0.0	c	0.0	c
	France	0.0	(0.0)	0.5	(0.1)	16.9	(0.6)	79.2	(0.6)	3.2	(0.2)	0.1	(0.0)	0.0	c
	Germany	0.4	(0.1)	8.1	(0.4)	46.4	(1.0)	44.0	(1.1)	1.1	(0.3)	0.0	(0.0)	0.0	c
	Greece	0.1	(0.0)	0.7	(0.2)	3.7	(0.5)	95.5	(0.6)	0.0	c	0.0	c	0.0	c
	Hungary	1.7	(0.3)	8.3	(0.5)	71.1	(0.7)	18.9	(0.6)	0.0	(0.0)	0.0	c	0.0	c
	Iceland	0.0	c	0.0	c	0.0	c	99.2	(0.1)	0.8	(0.1)	0.0	c	0.0	c
	Ireland	0.0	(0.0)	2.0	(0.2)	61.6	(0.7)	27.9	(0.9)	8.5	(0.7)	0.0	c	0.0	c
	Israel	0.0	(0.0)	0.1	(0.1)	16.7	(0.9)	82.4	(0.9)	0.7	(0.2)	0.0	(0.0)	0.0	c
	Italy	0.0	c	1.0	(0.2)	13.5	(0.5)	77.8	(0.5)	7.7	(0.3)	0.0	c	0.0	c
	Japan	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c	0.0	c
	Korea	0.0	c	0.0	c	16.1	(0.7)	83.8	(0.7)	0.1	(0.0)	0.0	c	0.0	c
	Latvia	0.7	(0.1)	9.8	(0.5)	86.0	(0.5)	2.5	(0.2)	0.0	(0.0)	0.0	c	1.1	(0.2)
	Lithuania	0.1	(0.1)	2.4	(0.2)	90.2	(0.5)	7.3	(0.4)	0.0	c	0.0	c	0.0	c
	Luxembourg	0.3	(0.1)	10.0	(0.1)	48.3	(0.1)	40.3	(0.1)	1.1	(0.1)	0.0	c	0.0	c
	Mexico	0.9	(0.2)	2.9	(0.4)	17.6	(1.1)	77.8	(1.0)	0.6	(0.1)	0.1	(0.1)	0.0	c
	Netherlands	0.1	(0.0)	2.6	(0.3)	36.8	(0.8)	59.3	(0.8)	1.2	(0.2)	0.0	(0.0)	0.0	c
	New Zealand	0.0	c	0.0	c	0.1	(0.0)	6.6	(0.5)	89.0	(0.4)	4.2	(0.2)	0.0	c
	Norway	0.0	c	0.0	c	0.3	(0.1)	99.3	(0.3)	0.4	(0.2)	0.0	c	0.0	c
	Poland	0.3	(0.1)	3.1	(0.3)	95.1	(0.5)	1.4	(0.4)	0.0	c	0.0	c	0.0	c
	Portugal	2.4	(0.2)	7.2	(0.4)	17.2	(0.9)	57.4	(1.3)	0.2	(0.1)	0.0	c	15.7	(1.5)
	Slovak Republic	1.9	(0.2)	4.3	(0.4)	40.8	(1.1)	51.3	(1.0)	1.7	(0.5)	0.0	c	0.0	c
	Slovenia	0.3	(0.0)	0.7	(0.2)	6.2	(0.4)	92.4	(0.4)	0.4	(0.1)	0.0	c	0.0	c
	Spain	0.0	(0.0)	5.9	(0.2)	24.1	(0.4)	69.9	(0.5)	0.1	(0.0)	0.0	c	0.0	c
	Sweden	0.0	c	2.1	(0.3)	96.3	(0.6)	1.6	(0.5)	0.0	c	0.0	c	0.0	c
	Switzerland	0.5	(0.1)	10.2	(0.6)	60.8	(1.4)	27.8	(1.4)	0.7	(0.3)	0.0	(0.0)	0.0	c
	Turkey	0.1	(0.1)	0.4	(0.2)	17.7	(1.1)	78.8	(1.1)	2.9	(0.3)	0.1	(0.0)	0.0	c
	United Kingdom	0.0	c	0.0	c	0.0	(0.0)	1.0	(0.6)	93.4	(0.6)	5.6	(0.2)	0.0	c
	United States	0.0	c	0.1	(0.1)	7.5	(0.5)	73.6	(0.8)	18.7	(0.7)	0.1	(0.1)	0.0	c

Note: The large number of students with missing grade-level information in Ukraine can be attributed to missing data from students in the first and second year of vocational colleges. Most of these 15-year-old students would have been in the first year of vocational college, which is equivalent to grade 10.

StatLink https://doi.org/10.1787/888934028862

Table I.A2.8 [2/2] **Percentage of students at each grade level**

		All students													
		7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above		Information unavailable	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners	Albania	0.2	(0.1)	1.2	(0.3)	36.6	(1.4)	61.5	(1.4)	0.5	(0.1)	0.0	(0.0)	0.0	c
	Argentina	2.1	(0.5)	9.8	(0.7)	22.1	(0.8)	63.8	(1.4)	1.8	(1.0)	0.0	(0.0)	0.4	(0.4)
	Baku (Azerbaijan)	0.2	(0.1)	2.8	(0.9)	34.7	(0.7)	61.5	(1.2)	0.7	(0.1)	0.0	c	0.0	c
	Belarus	0.1	(0.0)	0.9	(0.2)	42.8	(0.9)	56.2	(0.9)	0.0	c	0.0	c	0.0	c
	Bosnia and Herzegovina	0.0	(0.0)	0.2	(0.1)	16.2	(1.1)	83.4	(1.1)	0.1	(0.1)	0.0	c	0.0	c
	Brazil	4.1	(0.2)	8.1	(0.5)	13.5	(0.6)	33.5	(0.8)	39.3	(0.8)	1.5	(0.1)	0.0	c
	Brunei Darussalam	0.0	(0.0)	0.5	(0.1)	6.5	(0.1)	59.7	(0.1)	29.2	(0.1)	4.1	(0.0)	0.0	c
	B-S-J-Z (China)	0.3	(0.1)	1.5	(0.2)	38.7	(1.7)	58.2	(1.6)	1.3	(0.2)	0.0	(0.0)	0.0	c
	Bulgaria	0.2	(0.1)	2.7	(0.4)	92.8	(0.5)	4.2	(0.3)	0.0	(0.0)	0.0	c	0.0	c
	Costa Rica	4.8	(0.5)	13.8	(0.7)	36.5	(1.1)	44.7	(1.5)	0.2	(0.1)	0.0	c	0.0	c
	Croatia	0.0	(0.0)	0.3	(0.2)	78.9	(0.4)	20.8	(0.4)	0.0	c	0.0	c	0.0	c
	Cyprus	0.0	c	0.1	(0.1)	4.4	(0.4)	94.4	(0.4)	1.1	(0.1)	0.0	c	0.0	c
	Dominican Republic	6.4	(0.6)	12.5	(0.8)	23.6	(0.8)	43.8	(1.2)	12.6	(0.7)	1.2	(0.1)	0.0	c
	Georgia	0.1	(0.0)	0.5	(0.1)	14.3	(0.6)	84.2	(0.6)	1.0	(0.2)	0.0	c	0.0	c
	Hong Kong (China)	1.2	(0.2)	5.9	(0.5)	26.1	(0.9)	66.0	(1.1)	0.8	(0.5)	0.0	c	0.0	c
	Indonesia	3.4	(1.1)	8.1	(1.0)	33.7	(2.0)	49.2	(2.2)	4.2	(0.7)	1.4	(0.9)	0.0	c
	Jordan	0.2	(0.1)	1.6	(0.2)	11.2	(0.6)	87.0	(0.7)	0.0	c	0.0	c	0.0	c
	Kazakhstan	0.1	(0.0)	1.7	(0.1)	44.0	(0.7)	53.4	(0.7)	0.8	(0.1)	0.0	(0.0)	0.0	c
	Kosovo	0.0	c	0.4	(0.1)	23.2	(0.9)	74.6	(0.9)	1.7	(0.2)	0.0	(0.0)	0.0	c
	Lebanon	5.3	(0.5)	8.5	(0.5)	16.3	(0.9)	58.2	(1.0)	11.7	(0.5)	0.1	(0.1)	0.0	c
	Macao (China)	1.9	(0.1)	9.4	(0.2)	29.7	(0.2)	57.9	(0.2)	1.0	(0.1)	0.0	(0.0)	0.0	c
	Malaysia	0.0	c	0.0	c	5.5	(0.6)	94.2	(0.6)	0.3	(0.1)	0.0	c	0.0	c
	Malta	0.0	c	0.0	c	0.1	(0.0)	5.4	(0.2)	94.4	(0.1)	0.1	(0.0)	0.0	c
	Moldova	0.2	(0.1)	6.2	(0.5)	83.2	(0.8)	10.4	(0.8)	0.0	(0.0)	0.0	c	0.0	c
	Montenegro	0.0	c	0.0	c	3.3	(0.3)	93.8	(0.3)	2.9	(0.1)	0.0	c	0.0	c
	Morocco	8.0	(0.7)	13.9	(1.1)	32.1	(1.9)	38.4	(2.7)	7.7	(0.8)	0.0	c	0.0	c
	North Macedonia	0.0	c	0.2	(0.1)	95.8	(0.1)	4.0	(0.1)	0.0	c	0.0	c	0.0	c
	Panama	3.2	(0.5)	6.9	(0.6)	20.6	(1.0)	65.4	(1.4)	3.8	(0.4)	0.0	(0.0)	0.0	c
	Peru	1.8	(0.3)	5.7	(0.4)	14.3	(0.5)	54.5	(0.7)	23.6	(0.6)	0.0	c	0.0	c
	Philippines	4.5	(0.4)	12.8	(0.6)	51.1	(0.7)	30.9	(0.7)	0.6	(0.3)	0.0	(0.0)	0.0	c
	Qatar	1.3	(0.1)	4.5	(0.1)	18.0	(0.1)	63.4	(0.1)	12.9	(0.1)	0.0	(0.0)	0.0	c
	Romania	0.9	(0.3)	6.0	(0.9)	77.9	(0.9)	15.1	(0.5)	0.0	(0.0)	0.0	c	0.0	c
	Russia	0.4	(0.0)	7.7	(0.4)	81.1	(0.9)	10.7	(1.1)	0.1	(0.0)	0.0	c	0.0	c
	Saudi Arabia	1.2	(0.2)	3.6	(0.6)	14.0	(1.8)	77.5	(2.4)	3.6	(0.3)	0.1	(0.0)	0.0	c
	Serbia	0.1	(0.1)	0.8	(0.2)	87.7	(0.4)	11.4	(0.4)	0.0	c	0.0	c	0.0	c
	Singapore	0.0	(0.0)	1.1	(0.1)	7.6	(0.3)	90.8	(0.5)	0.4	(0.2)	0.0	c	0.0	c
	Chinese Taipei	0.0	c	0.1	(0.0)	35.7	(0.9)	64.2	(0.9)	0.0	(0.0)	0.0	c	0.0	c
	Thailand	0.2	(0.1)	0.7	(0.2)	19.9	(0.9)	76.6	(0.9)	2.5	(0.3)	0.0	c	0.0	c
	Ukraine	0.0	c	0.4	(0.1)	29.8	(1.3)	41.3	(1.8)	0.5	(0.1)	0.0	c	28.0	(2.4)
	United Arab Emirates	0.3	(0.1)	1.5	(0.1)	9.6	(0.3)	56.8	(0.6)	29.9	(0.5)	1.9	(0.2)	0.0	c
	Uruguay	4.2	(0.5)	11.2	(0.5)	20.5	(0.7)	63.4	(1.1)	0.6	(0.1)	0.0	c	0.0	c
	Viet Nam	0.2	(0.1)	0.8	(0.3)	4.0	(1.2)	92.3	(2.5)	0.0	(0.0)	0.0	c	2.7	(2.0)

Note: The large number of students with missing grade-level information in Ukraine can be attributed to missing data from students in the first and second year of vocational colleges. Most of these 15-year-old students would have been in the first year of vocational college, which is equivalent to grade 10.

StatLink https://doi.org/10.1787/888934028862

Tables available on line

https://doi.org/10.1787/888934028862

Notes

1. More precisely, PISA assessed students who were at least 15 years and 3 complete months old and who were at most 16 years and 3 complete months old (i.e. younger than 16 years, 2 months and roughly 30 days old), with a tolerance of one month on each side of this age window. If the PISA assessment was conducted in April 2018, as was the case in most countries, all students born in 2002 would have been eligible.

2. Educational institutions are generally referred to as schools in this publication, although some educational institutions (in particular, some types of vocational education establishments) may not be referred to as schools in certain countries.

3. As might be expected from this definition, the average age of students across OECD countries was 15 years and 9 months. The range in country means was 2 months and 13 days (0.20 year), from the minimum country mean of 15 years and 8 months to the maximum country mean of 15 years and 10 months (OECD, 2019[3]).

4. Such a comparison is complicated by first-generation immigrant students, who received part of their education in a country other than the one in which they were assessed. Mean scores in any country/economy should be interpreted in the context of student demographics within that country/economy.

5. Details for countries that applied different sampling designs are documented in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

6. Due to the small size of these education systems, all schools and all eligible students within these schools were included in the samples of Brunei Darussalam, Cyprus (see note 8), Iceland, Luxembourg, Macao (China), Malta, Montenegro and Qatar.

7. The threshold for an acceptable participation rate after replacement varies between 85% and 100%, depending on the participation rate before replacement.

8. In particular, in the case of the Netherlands and the United Kingdom, non-response bias analyses relied on direct measures of school performance external to PISA, typically from national assessments. More indirect correlates of school performance were analysed in Hong Kong (China) and the United States, due to the absence of national assessments. The non-response problem in Hong Kong (China) can be attributed to two causes: lack of initiative amongst schools and teachers to participate in PISA, and a large number of schools that were considered to be non-responding schools, as less than 50% of sampled students in these schools sat the assessment.

9. These exclusions refer only to those students with limited proficiency in the language of instruction/assessment. Exclusions related to the unavailability of test material in the language of instruction are not considered in this analysis.
10. The preliminary attribution of school codes in the process of selecting, and then excluding, students and schools may have resulted in the double exclusion (at both the school and student levels) of some of the students with special education needs in Sweden. As a result, the overall exclusion rate in Sweden may have been overestimated by (at most) 0.5 of a percentage point. In this scenario, the overall exclusion rate would still be over 10% and the highest amongst PISA-participating countries/economies.
11. The overall exclusion rate includes those students who were excluded at the school level (Column 6) and those students who were excluded within schools (Column 11); however, only students enrolled in non-excluded schools were affected by within-school exclusions, hence the presence of the term equivalent to 1 minus Column 6 (expressed as a decimal).
12. If the correlation between the propensity of exclusions and student performance were 0.3, then resulting mean scores would likely have been overestimated by 1 score point if the exclusion rate were 1%; by 3 score points if the exclusion rate were 5%; and by 6 score points if the exclusion rate were 10%. If the correlation between the propensity of exclusions and student performance were 0.5, then resulting mean scores would likely have been overestimated by 1 score point if the exclusion rate were 1%; by 5 score points if the exclusion rate were 5%; and by 10 score points if the exclusion rate were 10%. For this calculation, a model was used that assumed a bivariate normal distribution for performance and the propensity to participate.
13. Testing material was adapted to each country. Versions in the same language thus differed across countries, and students in Luxembourg who were not instructed in one of the three languages in which testing material was available (English, French and German) were unable to sit the PISA assessment, even if such material were available in their language of instruction in a different country.

References

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [3]

OECD (forthcoming), *PISA 2018 Results (Volume IV): Are Students Smart about Money?,* PISA, OECD Publishing, Paris. [2]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

ANNEX A3
Technical notes on analyses in this volume

STANDARD ERRORS, CONFIDENCE INTERVALS, SIGNIFICANCE TESTS AND P-VALUES

The statistics in this report represent estimates based on samples of students, rather than values that could be calculated if every student in every country had answered every question. Consequently, it is important to measure the degree of uncertainty of the estimates. In PISA, each estimate has an associated degree of uncertainty, which is expressed through a standard error. The use of confidence intervals provides a way to make inferences about the population parameters (e.g. means and proportions) in a manner that reflects the uncertainty associated with the sample estimates. If numerous different samples were drawn from the same population, according to the same procedures as the original sample, then in 95 out of 100 samples the calculated confidence interval would encompass the true population parameter. For many parameters, sample estimators follow a normal distribution and the 95% confidence interval can be constructed as the estimated parameter, plus or minus 1.96 times the associated standard error.

In many cases, readers are primarily interested in whether a given value in a particular country is different from a second value in the same or another country, e.g. whether girls in a country perform better than boys in the same country. In the tables and figures used in this report, differences are labelled as statistically significant when a difference of that size or larger, in either direction, would be observed less than 5% of the time in samples, if there were actually no difference in corresponding population values. Throughout the report, significance tests were undertaken to assess the statistical significance of the comparisons made.

Some analyses in this volume explicitly report *p*-values (e.g. Table I.B1.10). *p*-values represent the probability, under a specified model, that a statistical summary of the data would be equal to or more extreme than its observed value (Wasserstein, L. and Lazar, 2016[1]). For example, in Table I.B1.10, the *p*-value represents the likelihood of observing, in PISA samples, a trend equal to or more extreme (in either direction) than what is reported, when in fact the true trend for the country is flat (equal to 0).

RANGE OF RANKS (CONFIDENCE INTERVAL FOR RANKINGS OF COUNTRIES)

An estimate of the rank of a country mean, across all country means, can be derived from the estimates of the country means from student samples. However, because mean estimates have some degree of uncertainty, this uncertainty should also be reflected in the estimate of the rank. While mean estimates from samples follow a normal distribution, this is not the case of the rank estimates derived from these. Therefore, in order to construct a confidence interval for ranks, simulation methods were used.

Data are simulated assuming that alternative mean estimates for each relevant country follow a normal distribution around the estimated mean, with a standard deviation equal to the standard error of the mean. Some 10 000 simulations are carried out and, based on the alternative mean estimates in each of these simulations, 10 000 possible rankings for each country are produced. For each country, the counts for each rank are aggregated from the largest to smallest until they equal 9 750 or more. The range of ranks reported for each country includes all the ranks that have been aggregated (this procedure assumes unimodality of the distribution of rank estimates from samples, but makes no other assumption about this distribution). This means that the range-of-ranks estimates reported in Chapter 4 represent a 97.5% confidence interval for the rank statistic.

The main difference between the range of ranks (e.g. Table I.4.4) and the comparison of countries' mean performance (e.g. Table I.4.1) is that the former takes account of the multiple comparisons involved in ranking countries/economies, while the latter does not. Therefore, sometimes there is a slight difference between the range of ranks and counting the number of countries above a given country, based on pairwise comparisons of the selected countries' performance. For instance, OECD countries Australia, Denmark, Japan and the United Kingdom have similar mean performance and the same set of countries whose mean score is not statistically different from theirs, based on Table I.4.1; but the range of ranks amongst OECD countries for the United Kingdom and Japan can be restricted to be with 97.5% confidence between 7th and 15th, while the range of ranks for Australia and Denmark is narrower (between 8th and 14th for Australia; between 9th and 15th for Denmark) (Table I.4.4). When interest lies in examining countries' rankings, this range of ranks should be used.

The confidence level of 97.5% for the range-of-ranks estimate was chosen to limit paradoxical situations. Indeed, Tables I.4.1, I.4.2 and I.4.3 determine statistical significance using two-tailed tests, as is usual when testing for statistical significance of mean differences.

When interest lies in ranking two countries relative to each other, however, it is more appropriate to use one-tailed tests, as the procedure described above implicitly does. All cases where the mean score of country A ranks above the mean score of country B result in the same ranking between the two countries, regardless of how far A lies above B's mean score. For example, the estimate of the mean score of Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]") is higher than the estimate of the mean score of Singapore in reading, and the *p*-value for observing a difference of that size (or larger, but in the same direction) is 3.4%. In this situation, a two-tailed test for the difference in mean reading performance between B-S-J-Z (China) and Singapore cannot reject the null hypothesis of equal means at conventional levels of significance (the two-tailed 95%-confidence interval includes 0), but a one-tailed test would reject equality at the 95% level. When only two countries are involved in the comparison, a simple way of ensuring consistency between the range of ranks (one-tailed tests) and the comparison of countries' mean performance (two-tailed tests) is to set the confidence level for the confidence interval on rank statistics at 97.5%.

PARITY INDEX

The parity index for an indicator is used by the UNESCO Institute of Statistics to report on Target 4.5 of the Sustainable Development Goals. It is defined as the ratio of the indicator value for one group to the value for another group. Typically, the group more likely to be disadvantaged is in the numerator, and the parity index takes values between 0 and 1 (with 1 indicating perfect parity). However, in some cases the group in the numerator has a higher value on the indicator. To restrict the range of the parity index between 0 and 2, and to make its distribution symmetrical around 1, an adjusted parity index is defined in these cases.

For example, the gender parity index for the share of students reaching Level 2 proficiency on the PISA scale is computed from the share of boys (p_b) and the share of girls (p_g) reaching Level 2 proficiency as follows:

Equation I.A3.1

$$PI_{b,g} = \begin{cases} \frac{p_g}{p_b} & \text{if } p_b \geq p_g \\ 2 - \frac{p_b}{p_g} & \text{if } p_b < p_g \end{cases}$$

The "parity index" reported in Tables I.10.2 and I.B1.50 corresponds to the adjusted parity index as defined by the UNESCO Institute of Statistics (UNESCO Institute of Statistics, 2019[2]).

References

UNESCO Institute of Statistics (2019), *Adjusted parity index*, http://uis.unesco.org/en/glossary-term/adjusted-parity-index (accessed on 8 October 2019). [2]

Wasserstein, R. L. and **N. Lazar** (2016), "The ASA Statement on p-Values: Context, Process, and Purpose", *The American Statistician*, Vol. 70/2, pp. 129-133, http://dx.doi.org/10.1080/00031305.2016.1154108. [1]

ANNEX A4
Quality assurance

Quality assurance procedures were implemented in all parts of PISA 2018, as was done for all previous PISA surveys. The PISA 2018 Technical Standards (available on line at www.oecd.org/pisa) specify the way in which PISA must be implemented in each country, economy and adjudicated region. International contractors monitor the implementation in each of these and adjudicate on their adherence to the standards.

The consistent quality and linguistic equivalence of the PISA 2018 assessment instruments were facilitated by assessing the ease with which the original English version could be translated. Two source versions of the assessment instruments, in English and French, were prepared (except for the financial literacy assessment and the operational manuals, which were provided only in English) in order for countries to conduct a double translation design, i.e. two independent translations from the source language(s), and reconciliation by a third person. Detailed instructions for the localisation (adaptation, translation and validation) of the instruments for the field trial and for their review for the main survey, and translation/adaptation guidelines were supplied. An independent team of expert verifiers, appointed and trained by the PISA Consortium, verified each national version against the English and/or French source versions. These translators' mother tongue was the language of instruction in the country concerned, and the translators were knowledgeable about education systems. For further information on PISA translation procedures, see the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

The survey was implemented through standardised procedures. The PISA Consortium provided comprehensive manuals that explained the implementation of the survey, including precise instructions for the work of school co-ordinators and scripts for test administrators to use during the assessment sessions. Proposed adaptations to survey procedures, or proposed modifications to the assessment session script, were submitted to the PISA Consortium for approval prior to verification. The PISA Consortium then verified the national translation and adaptation of these manuals.

To establish the credibility of PISA as valid and unbiased and to encourage uniformity in conducting the assessment sessions, test administrators in participating countries were selected using the following criteria: it was required that the test administrator not be the reading, mathematics or science instructor of any student in the sessions he or she would conduct for PISA; and it was considered preferable that the test administrator not be a member of the staff of any school in the PISA sample. Participating countries organised an in-person training session for test administrators.

Participating countries and economies were required to ensure that test administrators worked with the school co-ordinator to prepare the assessment session, including reviewing and updating the Student Tracking Form; completing the Session Attendance Form, which is designed to record students' attendance and instruments allocation; completing the Session Report Form, which is designed to summarise session times, any disturbance to the session, etc.; ensuring that the number of test booklets and questionnaires collected from students tallied with the number sent to the school (for countries using the paper-based assessment) or ensuring that the number of USB sticks or external laptops used for the assessment were accounted for (for countries using the computer-based assessment); and sending or uploading the school questionnaire, student questionnaires, parent and teacher questionnaires (if applicable), and all test materials (both completed and not completed) to the national centre after the assessment.

The PISA Consortium responsible for overseeing survey operations implemented all phases of the PISA Quality Monitor (PQM) process: interviewing and hiring PQM candidates in each of the countries, organising their training, selecting the schools to visit, and collecting information from the PQM visits. PQMs are independent contractors located in participating countries who are hired by the international survey operations contractor. They visit a sample of schools to observe test administration and to record the implementation of the documented field-operations procedures in the main survey.

Typically, two or four PQMs were hired for each country, and they visited an average of 15 schools in each country. If there were adjudicated regions in a country, it was usually necessary to hire additional PQMs, as a minimum of five schools were observed in adjudicated regions.

Approximately one-third of test items are open-ended items in PISA. Reliable human coding is critical for ensuring the validity of assessment results within a country, as well as the comparability of assessment results across countries. Coder reliability in PISA 2018 was evaluated and reported at both within- and across-country levels. The evaluation of coder reliability was made possible by the design of multiple coding: a portion or all of the responses from each human-coded constructed-response item were coded by at least two human coders.

All quality-assurance data collected throughout the PISA 2018 assessment were entered and collated in a central data-adjudication database on the quality of field operations, printing, translation, school and student sampling, and coding. Comprehensive reports were then generated for the PISA Adjudication Group. This group was formed by the Technical Advisory Group and the Sampling Referee. Its role is to review the adjudication database and reports in order to recommend adequate treatment to preserve the quality of PISA data. For further information, see the *PISA 2018 Technical Report* (OECD, forthcoming[1]). Overall, the review suggested good adherence of national implementations of PISA to the technical standards. Despite the overall high quality of data, a few countries' data failed to meet critical standards or presented inexplicable anomalies, such that the Adjudication Group recommends a special treatment of these data in databases and/or reporting.

The major issues for adjudication discussed at the adjudication meeting are listed below:

- In Viet Nam, while no major standard violation was identified, there were several minor violations and the adjudication group has identified technical issues affecting the comparability of their data, an essential dimension of data quality in PISA. Viet Nam's cognitive data show poor fit to the item-response-theory model, with more significant misfit than any other country/ language group. In particular, selected-response questions, as a group, appeared to be significantly easier for students in Viet Nam than expected, given the usual relationship between open-ended and selected-response questions reflected in the international model parameters. In addition, for several selected-response items, response patterns are not consistent across field trial and main survey administrations, ruling out possible explanations of misfit in terms of familiarity, curriculum or cultural differences. For this reason, the OECD cannot currently assure full international comparability of the results.
- The Netherlands missed the standard for overall exclusions by a small margin. At the same time, in the Netherlands UH booklets, intended for students with special education needs, were assigned to about 17% of the non-excluded students. Because UH booklets do not cover the domain of financial literacy, the effective exclusion rate for the financial literacy additional sample is above 20%.The fact that students that receive support for learning in school were systematically excluded from the financial literacy sample results in a strong upward bias for the country mean and other population statistics. Therefore, the Netherlands' results in financial literacy may not be comparable to those of other counties or to results for the Netherlands from previous years. The Netherlands also missed the school response rate (before replacement) by a large margin, and could only reach close to an acceptable response rate through the use of replacement schools. Based on evidence provided in a non-response bias analysis, the Netherlands' results in reading, mathematics and science were accepted as largely comparable, but, in consideration of the low response rate amongst originally sampled schools, are reported with an annotation.
- Portugal did not meet the student-response rate standard. In Portugal, response rates dropped between 2015 and 2018. A student-non-response-bias analysis was submitted, investigating bias amongst students in grades 9 and above. Students in grades 7 and 8 represented about 11% of the total sample, but 20% of the non-respondents. A comparison of the linked responding and non-responding cases, using sampling weights, revealed that non-respondents tended to score about one-third of a standard deviation below respondents on the national mathematics examination (implying a "raw" upward bias of about 10% of a standard deviation on population statistics that are based on respondents only). At the same time, a significant proportion of the performance differences could be accounted for by variables considered in non-response adjustments (including grade level). Nevertheless, a residual upward bias in population statistics remained, even when using non-response adjusted weights. The non-response bias analysis therefore implies a small upward bias for PISA 2018 performance results in Portugal. The Adjudication Group also considered that trend comparisons and performance comparisons with other countries may not be particularly affected, because an upward bias of that size cannot be excluded even in countries that met the response-rate standard or for previous cycles of PISA. Therefore, Portugal's results are reported with an annotation.

While the adjudication group did not consider the violation of response-rate standards by Hong Kong (China) and the United States (see Annex A2) as major adjudication issues, they noted several limitations in the data used in non-response-bias analyses submitted by Hong Kong (China) and the United States. In consideration of the lower response rates, compared to other countries, the data for Hong Kong (China) and the United States are reported with an annotation.

In Spain, while no major standard violation was identified, subsequent data analyses identified sub-optimal response behaviours of some students. This was especially evident in the reading-fluency items. The reporting of Spain's reading performance will be deferred as this issue will be further investigated. For more details, see Annex A9.

Reference

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

ANNEX A5
How comparable are the PISA 2018 computer- and paper-based tests?

In the vast majority of participating countries, PISA 2018 was a computer-based assessment. However, nine countries – Argentina, Jordan, Lebanon, the Republic of Moldova, the Republic of North Macedonia, Romania, Saudi Arabia, Ukraine and Viet Nam – assessed their students' knowledge and skills in PISA 2018 using paper-based instruments. These paper-based tests were offered to countries that were not ready, or did not have the resources, to transition to a computer-based assessment.[1] The paper-based tests comprise a subset of the tasks included in the computer-based version of the tests, all of which were developed in earlier cycles of PISA according to procedures similar to those described in Chapter 2. No task that was newly developed for PISA 2015 or PISA 2018 was included in the paper-based instruments; consequently, the new aspects of the science and reading frameworks were not reflected in the paper-based tests.

This annex describes the differences between paper- and computer-based instruments, and what they imply for the interpretation of results.

DIFFERENCES IN TEST ADMINISTRATION AND CONSTRUCT COVERAGE

Over the past decades, digital technologies have fundamentally transformed the ways we read and manage information. Digital technologies are also transforming teaching and learning, and how schools assess students. To reflect how students and societies now commonly access, use and communicate information, starting with the 2015 assessment cycle, the PISA test was delivered mainly on computers. Existing tasks were adapted for delivery on screen; new tasks (initially only in science, then, for PISA 2018, also in reading) were developed that made use of the affordances of computer-based testing and that reflected the new situations in which students apply their science or reading skills in real life.

Because pen-and-paper tests are composed only of items initially developed for cycles up to PISA 2012, the paper-based version of the PISA 2018 test does not reflect the updates made to the assessment frameworks and to the instruments for science and reading. In contrast, the paper-based instruments for mathematics and their corresponding computer-based versions have their roots in the same framework, originally developed for PISA 2012.

The changes introduced in the assessment of science, in 2015, and of reading, in 2018, have deep implications for the set of assessment tasks used. The new frameworks resulted in a larger amount of assessment tasks at all levels; extended coverage of the reading and science scales through tasks that assess basic reading processes and emerging science skills (proficiency Levels 1b in science and 1c in reading); an expanded range of skills measured by PISA; and the inclusion of new processes or new situations in which students' competence manifests itself. Table I.A5.1 summarises the differences between the paper- and computer-based tests of reading; Table I.A5.2 summarises the corresponding differences in science.[2]

In reading, newly developed tasks could include using hyperlinks or other navigation tools (e.g. menus, scroll bars) to move between text segments. At the beginning of the reading test, a section was added to measure reading fluency, using timed sentence-comprehension tasks (see Chapter 1, Annex A6 and Annex C). None of these tasks would be feasible in a large-scale paper-based assessment. In science, new "interactive" tasks were developed for the PISA 2015 assessment. These tasks used computer simulations to assess students' ability to conduct scientific enquiry and interpret the resulting evidence. In these tasks, the information that students see on the screen is determined, in part, by their own interactions (through mouse clicks, keyboard strokes, etc.) with the task.

There are other differences between the PISA paper- and computer-based tests in addition to the tasks included in the tests and the medium through which students interacted with those tasks.

While the total testing time for all students was two hours, students who sat the test using computers had to take a break before starting work on the second half of the test, and had to wait until the end of the first hour before doing so. Students who sat the paper-based test also had to take a break after one hour of testing, but they could start working on the second half of the test during that first hour.

Another difference in test administration was that students who sat the test using computers could not go back to questions in a previous test unit or revise their answers during the test or after reaching the end of the test sequence (neither at the end of the first hour, nor at the end of the second hour).[3] In contrast, students who sat the paper-based version could, if they finished earlier, return to their unsolved tasks or change the answers they had originally given to some of the questions.

In 2018, and on average across countries that delivered the test on computer, 50% of students completed the reading test within about 40 minutes, i.e. about 20 minutes before the end of the test hour (Table I.A8.15). For additional analyses on response-time data, see Annex A8 and in the *PISA 2018 Technical Report* (OECD, forthcoming[1]).

In addition, the computer-based test in reading was a multi-stage adaptive test (see Chapter 1). In practice, the test forms consisted of three segments (stages): students were presented with a particular sequence of test tasks in the second and third stages based on a stochastic algorithm that took into account their performance on previous segments (OECD, forthcoming[1]; Yamamoto, Shin and Khorramdel, 2018[2]).[4] In science and mathematics (and also in reading for those countries that delivered the paper-based test), students were assigned test forms via a random draw, independent of the student's proficiency or behaviour on the test.

Table I.A5.1 **Differences between paper- and computer-based assessments of reading**

	Paper ("A" booklets)	Paper ("B" booklets)	Computer	Computer (excluding reading-fluency tasks)
Number of assessment tasks	88	87	309*	244*
Number of unique test booklets/forms	12	12	2304 possible paths through the assessment (12 reading fluency combinations x 192 adaptive paths)	192 possible paths through the assessment (128 unique combinations of items, of which 64 exist, in different disposition, as part of two paths)
Assignment of test booklets/forms to student	Random	Random	Random (reading fluency) + Adaptive	Adaptive
Assessment tasks, by PISA cycle in which they were first used				
PISA 2018	0	0	237	172
PISA 2009	49	59	44	44
PISA 2000	39	28	28	28
Range of task difficulty, on the PISA reading scale (RP62)**				
Min	224	224	67	224
10th percentile	377	373	213	378
Median	477	468	451	480
90th percentile	632	633	631	642
Max	1 045	1 045	1 045	1 045
Number of tasks, by proficiency level				
Level 6	4	4	10	10
Level 5	5	5	23	23
Level 4	18	14	34	34
Level 3	16	16	50	50
Level 2	22	23	71	71
Level 1a	20	18	47	46
Level 1b	2	5	31	9
Level 1c	1	2	12	1
Below Level 1c	0	0	31	0
Number of tasks, by sources required				
Single source	86	85	257	192
Multiple source	2	2	52	52
Number of tasks, by process				
Reading fluency	0	0	65	0
Locating information	17	18	49	49
Understanding	50	45	131	131
Evaluating and reflecting	21	24	64	64

Notes: "A" and "B" booklets have 72 test items in common; items unique to "A" booklets tend to be, on average, more difficult than the items unique to "B" booklets. Only items common to both "A" and "B" booklets are also used in the computer-based test. In the absence of adaptive testing, countries were invited to choose the booklet set that best matched the expected proficiency of their students. In 2018, only Ukraine used the "A" booklets; all other countries that delivered PISA 2018 on paper used the "B" booklets.

*Item CR563Q12, which was included in the computer-based test of reading but excluded from scaling, is not included in item counts in this table.

**All percentiles are unweighted. For the computer-adaptive test, the actual distribution of task difficulty, weighted by the proportion of students who responded to each task, is also a function of the distribution of student proficiency in the country.

Source: OECD, PISA 2018 Database; *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Table I.A5.2 **Differences between paper- and computer-based assessments of science**

	Paper	Computer
Number of assessment tasks	85	115
Number of unique test booklets/forms	18	18
Assignment of test booklets/forms to student	Random	Random
Assessment tasks, by PISA cycle in which they were first used		
PISA 2018	0	76
PISA 2006	85	39
Range of task difficulty, on the PISA reading scale (RP62)		
Min	305	305
10th percentile	437	426
Median	539	535
90th percentile	649	659
Max	821	925
Number of tasks, by proficiency level		
Level 6	3	3
Level 5	8	16
Level 4	23	30
Level 3	31	37
Level 2	16	21
Level 1a	3	7
Level 1b	1	1
Below Level 1b	0	0
Number of tasks, by science competency		
Interpret data and evidence scientifically	28	36
Explain phenomena scientifically	41	49
Evaluate and design scientific enquiry	16	30
Number of tasks, by type of knowledge		
Content	51	49
Procedural	24	47
Epistemic	10	19
Number of tasks, by system		
Living	39	47
Earth and Space	18	30
Physical	28	38

Source: OECD, PISA 2018 Database; *PISA 2018 Technical Report* (OECD, forthcoming[1]).

HOW THE EVIDENCE ABOUT MODE EFFECTS WAS USED TO LINK THE TWO DELIVERY FORMATS

In order to ensure comparability of results between the computer-delivered tasks and the paper-based tasks that were used in previous PISA assessments (and are still in use in countries that use paper instruments), for the test items common to the two administration modes, the invariance of item characteristics was investigated using statistical procedures. These included model-fit indices to identify measurement invariance (see Annex A6), and a randomised mode-effect study in the PISA 2015 field trial that compared students' responses to paper-based and computer-delivered versions of the same tasks across equivalent international samples (OECD, 2016[3]). For the majority of items, the results supported the use of common difficulty and discrimination parameters across the two modes of assessment. For some items, however, the computer-delivered version was found to have a different relationship with student proficiency from the corresponding, original paper version. Such tasks had different difficulty parameters (and sometimes different discrimination parameters) in countries that delivered the test on computer. In effect, this partial invariance approach both accounts for and corrects the potential effect of mode differences on test scores.

Table I.A5.3 shows the number of anchor items that support the reporting of results from the computer-based and paper-based assessments on a common scale. The large number of items with common difficulty and discrimination parameters indicates a strong link between the scales. This strong link corroborates the validity of mean comparisons across countries that delivered the test in different modes. At the same time, Table I.A5.3 also shows that a large number of items used in the PISA 2018 computer-based tests of reading and, to a lesser extent, science, were not delivered on paper. Caution is therefore required when drawing conclusions about the meaning of scale scores from paper-based tests, when the evidence that supports these conclusions is based on the full set of items. For example, the proficiency of students who sat the PISA 2018 paper-based test of reading should be described in terms of the PISA 2009 proficiency levels, not the PISA 2018 proficiency levels, and similarly for science. This means, for example, that even though PISA 2018 developed a description of the skills of students who scored below Level 1b in reading, it remains unclear whether students who scored within the range of Level 1c on the paper-based tests have acquired these basic reading skills.

Table I.A5.3 **Anchor items across paper- and computer-based scales**

Scalar-invariant, metric-invariant and unique items in PISA 2018 paper and computer tests

	Reading	Mathematics	Science
Items with common difficulty and discrimination parameters across modes (scalar invariant)	40	50	29
Items with common discrimination parameter across modes, but distinct difficulty parameter (metric invariant)	32	31	10
Items with mode-specific parameters	0	1*	0
Items not delivered on computer (paper-based only)	15 (“A” booklets) 16 (“B” booklets)	1	46
Items not delivered on paper (computer-based only)	172+65 “fluency” items**	0	76

* In PISA 2015 and in the mode-effect study, Item M192Q01 was excluded from scaling in the computer-based version due to a technical issue. Its parameters were therefore freely estimated in 2018.

** In addition, item CR563Q12 was included in the computer-based test of reading but excluded from scaling due to a technical problem with the recording of students' answers.

Note: The table reports the number of scalar-invariant, metric-invariant and unique items based on international parameters. In any particular country, items that receive country-specific item parameters (see Annex A6) must also be considered.

Source: OECD, PISA 2018 Database; *PISA 2018 Technical Report* (OECD, forthcoming[1]).

Notes

1. Albania, Georgia, Indonesia, Kazakhstan, Kosovo, Malta, Panama and Serbia transitioned to the computer-based assessment in 2018. All other returning PISA 2018 participants, including all OECD countries, made the transition in 2015.
2. No subscales are estimated for students who sat the paper-based test of reading.
3. In the computer-based test, and with limited exceptions, students were still able to go back to a previous question within the same unit and revisit their answers. They were not allowed to go back to a previous unit.
4. Before the first segment of the adaptive test (also called “core” stage), all students also completed a 3-minute reading-fluency section, which consisted of 21 or 22 items per student, assembled, according to 12 possible combinations, from 65 available items. Performance on this reading-fluency section was not considered by the adaptive algorithm in the main section of the reading test.

References

OECD (2016), *The PISA 2015 Field Trial Mode-Effect Study*, OECD Publishing, Paris, www.oecd.org/pisa/data/PISA-2015-Vol1-Annex-A6-PISA-2015-Field-Trial-Mode-Effect-Analysis.pdf (accessed on 1 July 2019). [3]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

Yamamoto, K., H. Shin and **L. Khorramdel** (2018), “Multistage Adaptive Testing Design in International Large-Scale Assessments”, *Educational Measurement: Issues and Practice*, Vol. 37/4, pp. 16-27, http://dx.doi.org/10.1111/emip.12226. [2]

ANNEX A6
Are PISA reading scores comparable across countries and languages?

The validity and reliability of PISA scores, and their comparability across countries and languages, are the key concerns that guide the development of the assessment instruments and the selection of the statistical model for scaling students' responses. The procedures used by PISA to meet these goals include qualitative reviews conducted by national experts on the final main study items and statistical analyses of model fit in the context of multi-group item-response-theory models, which indicate the measurement equivalence of each item across groups defined by country and language.

COUNTRIES' PREFERRED ITEMS

National reading experts conducted qualitative reviews of the full set of items included in the PISA 2018 assessment at different stages of their development. The ratings and comments submitted by national experts determined the revision of items and coding guides for the main study, and guided the final selection of the item pool. In many cases, these changes mitigated cultural concerns and improved test fairness. At the end of 2018, the PISA consortium asked national experts to confirm or revise their original ratings, with respect to the final instruments. Sixty-five national centres submitted ratings of the relevance of PISA 2018 reading items to measure students' "preparedness for life" – a key aspect of the validity of PISA (response options were: "not at all relevant", "somewhat relevant", "highly relevant"). National experts also indicated whether the specific competences addressed by each item were within the scope of official curricula ("not in curriculum", "in some curricula", "standard curriculum material"). While PISA does not intend to measure only what students learn as part of the school curriculum, ratings of curriculum coverage for PISA items provide contextual indicators to understand countries' strengths and weaknesses in the assessment.

On average across countries/economies, 76% of items were rated as "highly relevant for students' preparedness for life" (the highest possible rating); only 3% received a low rating on this dimension (rating equal to 1). Thirty-five out of 65 countries/economies did not rate any item as being "not relevant" to students' preparedness for life.

On the other hand, many national experts indicated less overlap between national curricula and the PISA reading item set. On average, 63% of items were rated as "standard curriculum material", and 9% of items were identified as "not in curriculum". National experts from six countries – Australia, Costa Rica, Estonia, Finland, Iceland and the Republic of Moldova – indicated that all items used in PISA could be considered standard curriculum material in their country.

Table I.A6.1 provides a summary of the ratings received from national centres about the PISA 2018 set of reading items.

NATIONAL ITEM DELETIONS, ITEM MISFIT, AND ITEM-BY-COUNTRY INTERACTIONS

PISA reporting scales in reading, mathematics and science are linked across countries, survey cycles and delivery modes (paper and computer) through common items whose parameters are constrained to the same values and which can therefore serve as "anchors" on the reporting scale. A large number of anchor items support the validity of cross-country comparisons and trend comparisons.

The unidimensional multi-group item-response-theory models used in PISA, with groups defined by language within countries and by cycle, also result in model-fit indices for each item-group combination. These indices can indicate tensions between model constraints and response data, a situation known as "misfit" or "differential item functioning" (DIF).

In cases where the international parameters for a given item did not fit well for a particular country or language group, or for a subset of countries or language groups, PISA allowed for a "partial invariance" solution, in which the equality constraints on the item parameters were released and group-specific item parameters were estimated. This approach was favoured over dropping the group-specific item responses for these items from the analysis in order to retain the information from these responses. While the items with DIF, treated in this way, no longer contribute to the international set of comparable responses, they help reduce measurement uncertainty for the specific country-by-language group.

In rare instances where the partial invariance model was not sufficient to resolve the tension between students' responses and the IRT model, the group-specific response data for that particular item were dropped.

An overview of the number of international/common (invariant) item parameters and group-specific item parameters in reading for PISA 2018 is given in Figure I.A6.1 and Figure I.A6.2; the corresponding figures for other domains can be found in the *PISA 2018 Technical Report* (OECD, forthcoming[1]). Each set of stacked bars in these figures represents a country or economy (for countries and economies with multiple language groups, a weighted average of the scaling groups is presented).

Table I.A6.1 [1/2] **How national experts rated PISA reading items**

Percentage of test items, by rating

		In curriculum?			Relevant to "preparedness for life"?		
		Not in curriculum (%)	In some curricula (%)	Standard curriculum material (%)	Not at all relevant (%)	Somewhat relevant (%)	Highly relevant (%)
OECD	Australia	0.0	0.0	100.0	0.0	0.0	100.0
	Austria	0.4	20.0	79.6	2.0	33.9	64.1
	Belgium (Flemish Community)	0.0	9.0	91.0	0.0	2.0	98.0
	Belgium (French Community)	0.4	5.0	94.6	0.0	5.0	95.0
	Canada	0.0	26.9	73.1	0.0	15.9	84.1
	Chile	0.8	28.6	70.6	5.3	14.3	80.4
	Colombia	1.3	14.4	84.3	1.3	3.4	95.3
	Czech Republic	2.9	45.7	51.4	0.4	39.2	60.4
	Denmark	0.0	45.7	54.3	0.0	29.8	70.2
	Estonia	0.0	0.0	100.0	0.0	0.0	100.0
	Finland	0.0	0.0	100.0	0.0	0.0	100.0
	France	22.9	28.6	48.6	3.7	14.3	82.0
	Germany	0.0	9.0	91.0	0.0	0.8	99.2
	Greece	9.0	28.6	62.4	4.9	2.0	93.1
	Hungary	20.4	52.7	26.9	0.0	23.7	76.3
	Iceland	0.0	0.0	100.0	0.0	3.7	96.3
	Israel	10.2	26.1	63.7	9.0	44.5	46.5
	Italy	5.3	28.3	66.4	5.7	4.1	90.2
	Japan	1.2	0.4	98.4	1.2	0.4	98.4
	Korea	0.0	13.1	86.9	0.0	0.4	99.6
	Latvia	0.0	7.8	92.2	0.0	3.7	96.3
	Luxembourg	0.0	11.8	88.2	0.0	0.0	100.0
	Mexico	0.0	15.7	84.3	0.0	0.0	100.0
	Netherlands	0.8	46.5	52.7	0.0	14.7	85.3
	New Zealand	0.0	18.8	81.2	0.0	11.4	88.6
	Norway	8.6	14.3	77.1	6.5	5.7	87.8
	Poland	0.4	14.3	85.3	0.0	0.8	99.2
	Portugal	53.9	24.1	22.0	20.0	31.0	49.0
	Slovak Republic	0.0	85.3	14.7	0.4	35.5	64.1
	Slovenia	27.3	20.0	52.7	8.2	46.5	45.3
	Sweden	0.8	19.7	79.5	0.0	11.6	88.4
	Switzerland	0.0	31.8	68.2	0.0	0.4	99.6
	United States	m	m	m	m	m	m

Note: Percentages may not add up to 100% due to rounding. Percentages are reported as a proportion of all test items that received a rating. For countries that delivered the test on paper, only ratings for trend items were considered. Countries and economies that are not included in this table did not submit ratings on the final set of items. In Switzerland, three experts from distinct language regions reviewed the items. For the few items where their ratings differed, a national rating was determined as follows: for relevance to "preparedness for life", the modal rating was considered; for curriculum overlap, the rating "in some curricula" was used unless all three experts agreed on one of the two other options. For Belgium, ratings are reported separately for the Flemish Community and for the French Community. For Denmark, the category "in some curricula" should be interpreted as "partly relevant to" the (single) national learning standards. Ratings for the United States are reported as missing; the education system in the United States is highly decentralised, with over 13 600 school districts that make curriculum decisions based on state recommendations. This makes it difficult to determine curriculum coverage in relation to assessment items.

StatLink https://doi.org/10.1787/888934028881

The bars represent the items used in the country. A colour-code indicates whether international item parameters were used in scaling (the same as in PISA 2015), or whether, due to misfit when using international parameters, national item parameters were used.[1] For items where international equality constraints were released, a distinction is made between two groups:

- items that received unique parameters for the particular group defined by country/language and year (in many cases, equality constraints across a subset of misfit groups defined by country/language and year, e.g. across all language groups in a country, could be implemented)
- items for which the "non-invariant" item parameters used in 2018 could be constrained to the same values used in 2015 for the particular country/language group (these items contribute to measurement invariance over time, but not across groups).

Table I.A6.1 [2/2] **How national experts rated PISA reading items**

Percentage of test items, by rating

		In curriculum?			Relevant to "preparedness for life"?		
		Not in curriculum (%)	In some curricula (%)	Standard curriculum material (%)	Not at all relevant (%)	Somewhat relevant (%)	Highly relevant (%)
Partners	**Albania**	23.7	19.2	57.1	11.0	31.8	57.1
	Argentina	26.4	20.8	52.8	12.5	19.4	68.1
	Baku (Azerbaijan)	0.4	96.7	2.9	0.0	10.7	89.3
	Belarus	0.0	13.1	86.9	0.0	41.2	58.8
	Brazil	0.0	3.7	96.3	1.2	4.1	94.7
	Brunei Darussalam	21.2	63.3	15.5	22.4	58.0	19.6
	B-S-J-Z (China)	1.2	13.1	85.7	0.4	6.1	93.5
	Bulgaria	0.0	22.9	77.1	0.0	31.0	69.0
	Costa Rica	0.0	0.0	100.0	0.0	0.0	100.0
	Croatia	21.6	48.2	30.2	0.0	17.6	82.4
	Cyprus	0.0	33.9	66.1	0.0	5.7	94.3
	Hong Kong (China)	5.7	46.9	47.3	0.8	41.2	58.0
	Jordan	11.1	25.0	63.9	6.9	8.3	84.7
	Kazakhstan	0.0	82.9	17.1	0.0	29.8	70.2
	Macao (China)	58.8	41.2	0.0	20.8	70.6	8.6
	Malaysia	6.5	51.4	42.0	0.4	42.9	56.7
	Malta	2.4	40.4	57.1	0.4	49.0	50.6
	Moldova	0.0	0.0	100.0	2.8	5.6	91.7
	Montenegro	2.9	4.5	92.7	5.7	17.1	77.1
	Morocco	24.9	47.8	27.3	3.3	40.0	56.7
	Panama	0.0	59.2	40.8	0.0	95.5	4.5
	Peru	0.0	18.4	81.6	0.0	3.7	96.3
	Qatar	2.5	50.4	47.1	0.0	9.4	90.6
	Romania	0.0	5.6	94.4	1.4	6.9	91.7
	Russia	17.2	20.9	61.9	0.0	55.3	44.7
	Serbia	68.6	18.8	12.7	0.0	1.6	98.4
	Singapore	0.8	0.4	98.8	0.0	6.5	93.5
	Chinese Taipei	0.0	86.9	13.1	0.0	75.9	24.1
	Thailand	0.0	18.4	81.6	0.0	7.3	92.7
	Ukraine	18.1	11.1	70.8	0.0	1.4	98.6
	United Arab Emirates	46.1	18.8	35.1	14.7	43.3	42.0
	Uruguay	9.4	36.5	54.1	7.3	36.1	56.7
	Viet Nam	45.8	51.4	2.8	45.8	51.4	2.8

Note: Percentages may not add up to 100% due to rounding. Percentages are reported as a proportion of all test items that received a rating. For countries that delivered the test on paper, only ratings for trend items were considered. Countries and economies that are not included in this table did not submit ratings on the final set of items. In Switzerland, three experts from distinct language regions reviewed the items. For the few items where their ratings differed, a national rating was determined as follows: for relevance to "preparedness for life", the modal rating was considered; for curriculum overlap, the rating "in some curricula" was used unless all three experts agreed on one of the two other options. For Belgium, ratings are reported separately for the Flemish Community and for the French Community. For Denmark, the category "in some curricula" should be interpreted as "partly relevant to" the (single) national learning standards. Ratings for the United States are reported as missing; the education system in the United States is highly decentralised, with over 13 600 school districts that make curriculum decisions based on state recommendations. This makes it difficult to determine curriculum coverage in relation to assessment items.

StatLink https://doi.org/10.1787/888934028881

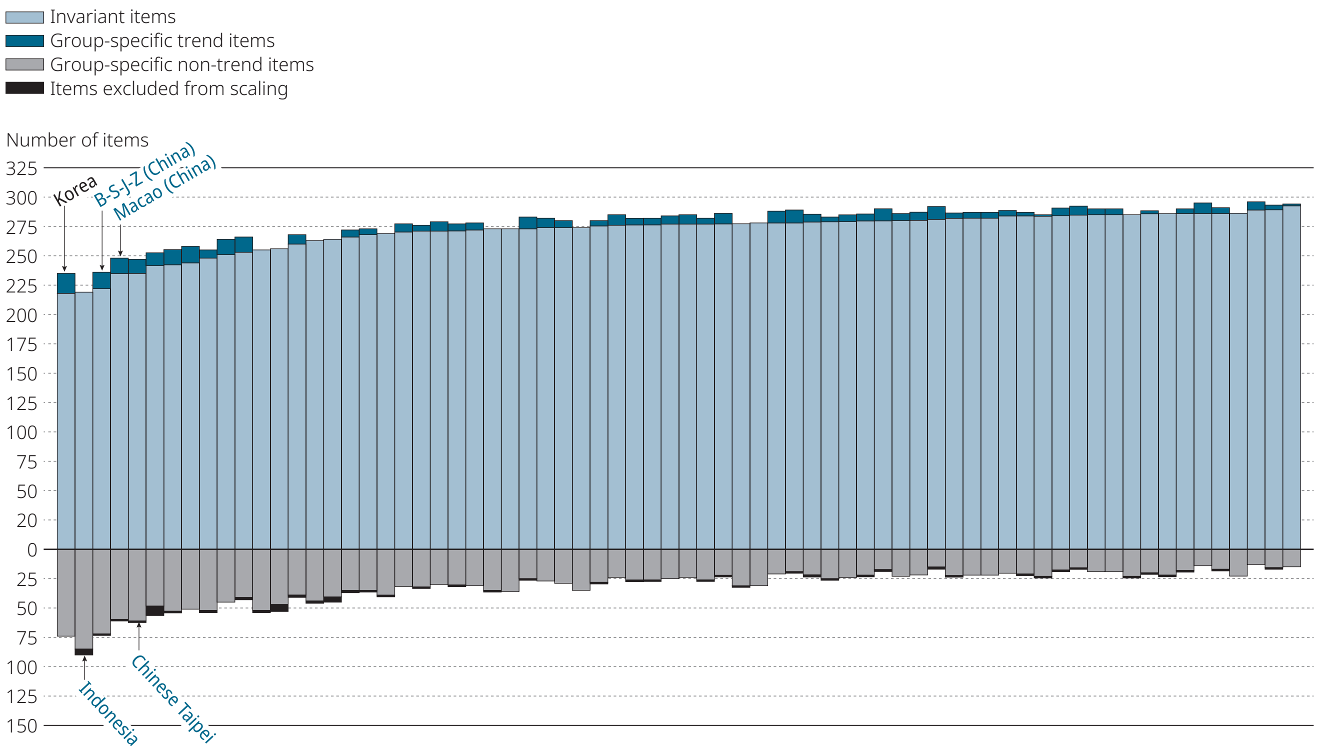
For any pair of countries/economies, the larger the number and share of common item parameters, the more comparable the PISA scores. As the figures show, comparisons between most countries' results are supported by strong links involving many items (in 58 of 79 countries/economies, over 85% of the items use international, invariant item parameters). Across every domain, international/common (invariant) item parameters dominate and only a small proportion of the item parameters are group-specific. The *PISA 2018 Technical Report* (OECD, forthcoming[1]) includes an overview of the number of deviations per item across all country-by-language groups.

The country/language group with the largest amount of misfit across items is Viet Nam (the same was found in mathematics and science too). The proportion of international trend items is between 50% and 60% in each subject. A similar level of misfit was also found in PISA 2015.

The possible reasons why the item-response theory model that fits all other countries well is not a good fit for Viet Nam's data are still being investigated. Initial analyses explored, at the item level, the direction of misfit (using mean deviation statistics), the characteristics of misfit items, and any potential sign of data manipulation or coder bias. For example, students' booklets were inspected, and the answers were compared to the codes included in the database. The analysis also involved comparisons of booklets and response patterns in PISA 2018 with the PISA 2015 main study and with the PISA 2015 and 2018 field trials.

Figure I.A6.1 **Invariance of items in the computer-based test of reading across countries/economies and over time**

Analyses based on 309 items (including reading-fluency tasks)



Notes: Each set of stacked columns corresponds to a distinct country/economy. For countries/economies with more than one scaling group, a weighted average of invariant and non-invariant items across scaling groups is reported.

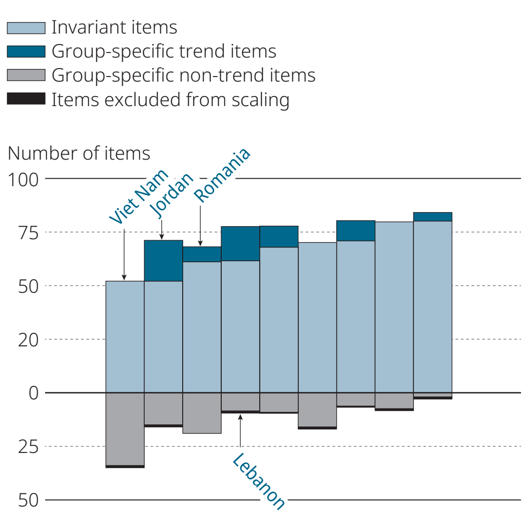
Item CR563Q12 was excluded from scaling in all countries and is not included among the 309 items considered for this figure.

Source: OECD, PISA 2018 Database; *PISA 2018 Technical Report* (OECD, forthcoming[1]).

StatLink https://doi.org/10.1787/888934028900

Figure I.A6.2 **Invariance of items in the paper-based test of reading across countries and over time**

Analyses based on 88 items ("A" booklets) or 87 items ("B" booklets)



Note: Each set of stacked columns corresponds to a distinct country. For countries with more than one scaling group, a weighted average of invariant and non-invariant items across scaling groups is reported.

Source: OECD, PISA 2018 Database; *PISA 2018 Technical Report* (OECD, forthcoming[1]).

StatLink https://doi.org/10.1787/888934028919

Indeed, while overall performance can vary across PISA administrations (and particularly between the field trial and the main study), the item-response patterns, conditional on overall performance, should remain relatively stable across administrations, unless the patterns are strongly influenced by test conditions, such as the print quality.

This initial investigation did not find any evidence of data manipulation or coder bias. Initial findings indicate that a significant amount of misfit could be modelled as a country-specific response-format effect, meaning that selected-response questions, as a group, appeared to be significantly easier for students in Viet Nam than expected, given the usual relationship between open-ended and selected-response questions reflected in the international model parameters. The initial investigation also found that for a number of selected-response items, response patterns were not consistent across field-trial and main study administrations. This inconsistency over time within the same country cannot be explained by familiarity, curriculum or cultural differences. After reviewing the data for Viet Nam, the PISA Adjudication Group concluded that targeted training and coaching on PISA-like items (and occasional errors induced by training or coaching) constitutes the most plausible explanation for the differences between student-response patterns observed in Viet Nam in 2018 and those observed in other countries or in previous cycles.

Whatever its causes, the statistical uniqueness of Viet Nam's response data implies that performance in Viet Nam cannot be validly reported on the same PISA scale as performance in other countries. It may still be possible to estimate an item-response-theory model for Viet Nam and report performance on a scale that retains some level of within-country trend comparability, but this scale could not be used to compare Viet Nam with other countries and could not be interpreted in terms of international proficiency levels.

In addition, Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]"), Indonesia, Korea, Macao (China) and Chinese Taipei (as well as, amongst countries that delivered the PISA test in the paper-based format, Jordan, Lebanon and Romania) show a relatively large number of patterns that are unexpected, based on international item parameters and given the overall performance level observed in these countries/economies. In all of these countries/economies, except Jordan, items with group-specific parameters and items excluded from scaling represent between 23% and 30% of all items in reading (in Jordan, they represent 40% of items in reading, 39% in science and 13% in mathematics). This mirrors earlier findings that differential item functioning in the PISA reading test is higher in Asian countries and in countries using non-Indoeuropean languages (Grisay and Monseur, 2007[2]; Grisay, Gonzalez and Monseur, 2009[3]). Another tentative pattern that can be established is that item misfit is higher, in reading, in countries where the language of assessment is not the language spoken outside of school by many of their students; this is the case in Indonesia and Lebanon. In these cases, the target construct for reading items may be confounded by language proficiency.

While the number of items affected is relatively large, the nature and extent of misfit are unlikely to affect the validity and comparability of PISA results in these cases. For example, in each of the countries/economies that delivered PISA on computer, including B-S-J-Z (China), Indonesia, Korea, Macao (China) and Chinese Taipei, comparisons of reading scores across countries are supported by at least 218 items with common, invariant parameters. In the case of Jordan and Lebanon, while international comparability is lower (also because these countries used paper-based instruments; see Annex A5), trend comparability is strong: for a majority of the items receiving country-specific item parameters, the observed response patterns are consistent with what had already been observed in 2015.

ARE PISA RANKINGS DETERMINED BY THE SELECTION OF ITEMS FOR THE TEST?

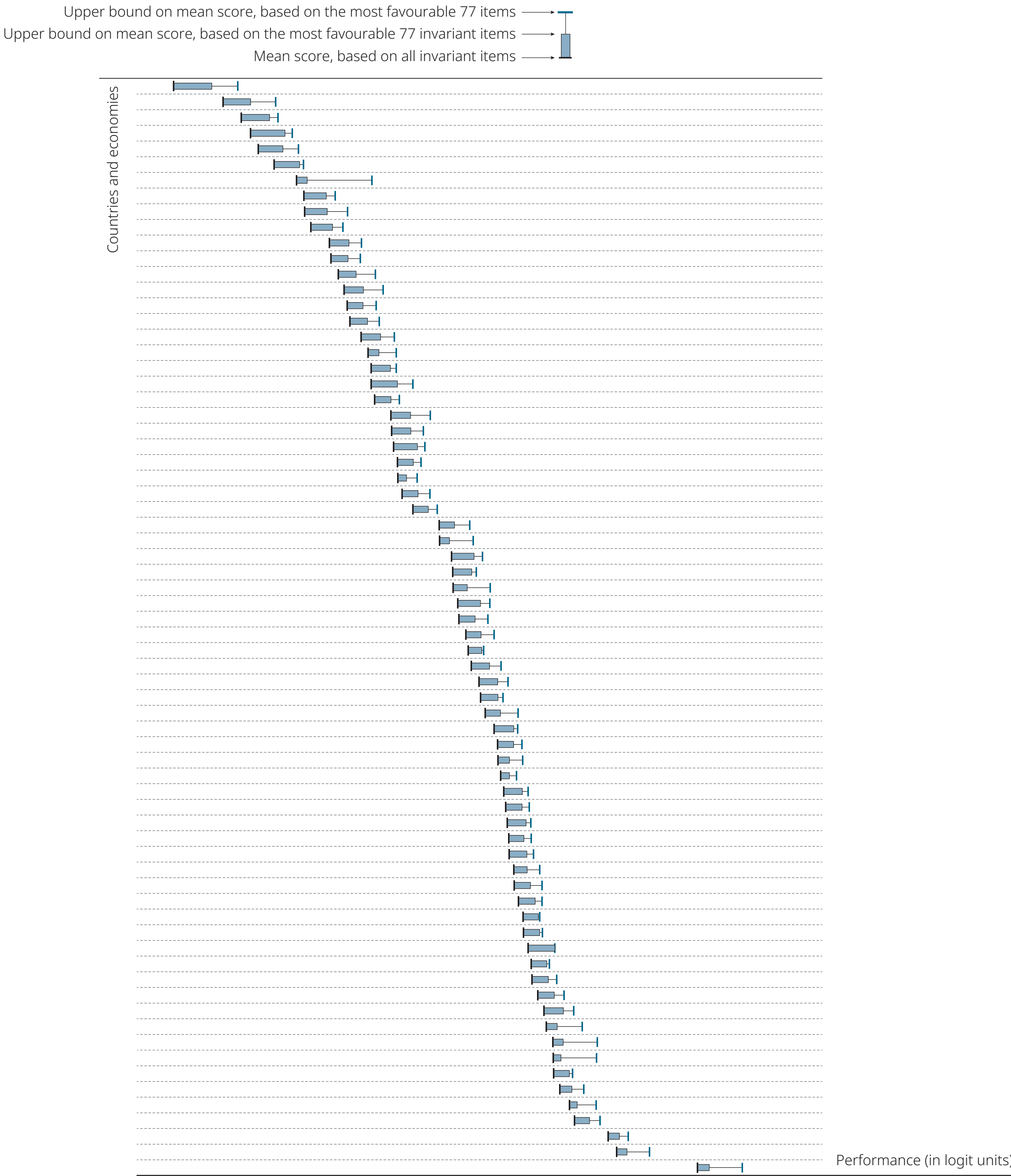
A key assumption of a fully invariant "international" item-response model is that a single model can describe the relationship between student proficiency and (international) item characteristics for all countries and economies. This would imply, for example, that any sufficiently large subset of items would result in the same performance estimate for the country/economy, up to a small "measurement error". In practice, the assumption of full invariance is relaxed in PISA, which estimates a "partial" invariance model that allows some items to have country/language-specific characteristics (see above). This strongly limits the impact of item selection on performance scores.

This section analyses the impact of item selection on mean-score rankings for countries that delivered the PISA 2018 test on computer. It does so both in the context of a hypothetical fully invariant item-response model and in the context of the partial-invariance model used in PISA. In both situations, the analysis asks: to what extent could a country improve its ranking simply through a more favourable selection of test items (i.e. without changing students' behaviour)?

In particular, for each country, three approximate measures of mean performance are computed: one based on the full set of invariant items, which is used as a reference, and two "upper bound" estimates based on more favourable sets of items. These upper bound estimates are based on two-thirds of the items only. In the "strong invariance" case, all items are considered when selecting the most-favourable 77 items (out of 115 items available in total); in the "partial invariance" case, only items that are scaled using international trend items are considered when selecting the most-favourable 77 items for each country/economy.

Figure I.A6.3 **Robustness of country mean scores in science**

Mean performance and upper bound on mean performance based on most favourable selection of 77 items



Note: Mean performance is computed based on invariant items only as the mean of logit-transformed percentages of correct answers, centred around the international mean and divided by the median absolute deviation. The value of 0 corresponds to the international mean for computer-based countries. To compute the upper bound on mean performance, only the most favourable 77 items (i.e. about two-thirds of the overall set of items) are considered for each country. The high mark selects these 77 items among all 115 items, assuming that they are invariant and can be used to compare countries; the more narrow range assumes that only the science items that are scaled with international item parameters are comparable across countries and economies.

Source: OECD, PISA 2018 Database.

StatLink https://doi.org/10.1787/888934028938

To avoid embedding other model assumptions in the comparison, country mean scores are not computed through an item-response model, but as simple averages of logit-transformed percent-correct statistics, centred around the international mean for each item.[2] The average score for a country whose students succeed at the international mean level on each item is therefore 0. Positive scores indicate that the country has, on average across items, higher success rates than the international mean; negative scores indicate that the country has, on average, lower success rates than the international mean.

The analysis in this section is based on the science test, because item-level statistics, including the percentage of correct answers or its logit-transformed values, are not directly comparable across countries for the reading test, which was delivered in adaptive fashion. The analysis intends to illustrate what the observed level of misfit implies for the substantive conclusions that are drawn from PISA, both before any country- and language-specific parameters are assigned, and after the set of invariant items is tailored to each country. Because the amount of model misfit is similar in every domain, the qualitative conclusions are expected to generalise to reading too.

The analysis shows that the selection of items only minimally affects the most important comparative conclusions – for example, whether a country scores above or below another country, on average – and that the influence of item selection on country rankings is reduced particularly when the partial-invariance model that PISA applies to student responses is duly considered. This means that the potential for improving a country's mean performance in PISA through a more favourable selection of items, indicated by the blue segments in Figure I.A6.3, is small in comparison to the overall variation in performance across countries.

ARE MEASURES OF READING FLUENCY COMPARABLE ACROSS COUNTRIES AND LANGUAGES?

Reading-fluency tasks required test-takers to decide as quickly as possible whether a simple sentence made sense (see Annex C).

Student scores on reading-fluency tasks (i.e. whether they correctly affirmed that a meaningful sentence made sense and rejected meaningless sentences) were considered together with the remaining reading tasks during scaling. These tasks amount to very simple literal understanding tasks. The analysis of country-by-item effects (DIF) did not highlight particular issues related to this group of items.

Timing information, however, was not used during scaling.[3] An initial analysis of item completion time for reading-fluency tasks indeed showed considerable country differences and, most important, item-by-country effects. For this reason, the Reading Expert Group that guided the development of the reading test does not recommend the use of time data at the item level as part of the international PISA reports, nor the construction of a simple international timing-based measure of fluency. At the same time, the Reading Expert Group supports the use of timing-based measures of fluency in national analyses, and encourages further research into the modelling of timing and accuracy data at national and international levels. Simple, descriptive measures of the total time spent by students on reading-fluency tasks are provided in Table I.A8.19 (available on line).

Data about response time and score (correct/incorrect) are available for all items, including reading fluency items, and for all students, as part of the public-use cognitive database. Interested researchers can access these data through the PISA website at www.oecd.org/pisa.

Notes

1. For countries that distributed the paper-based test, group invariance is assessed with respect to international paper-based item parameters. When comparing countries using the paper-based test to countries using the computer-based test, the number and share of items for which the difficulty parameter differs (metric invariant items; see Table I.A5.3) should also be considered.
2. The approximate mean scores used in Figure I.A6.3, based on logit-transformed and centred percent-correct statistics for invariant items, correlate at $r = 0.998$ (N = 70) with the mean scores based on plausible values reported in Table I.B1.6.
3. Timing information is collected and reported in databases for all items in the computer-based test, but is not considered, in general, part of the construct that is being assessed by these items. In contrast, in the case of reading-fluency items, both "speed" and "accuracy" are important aspects of the target construct, and students were explicitly told that their completion time would be considered, along with their answers ("You will have three minutes to read and respond to as many sentences as you can"). For this reason, the question whether timing information should be included in scaling was considered.

References

Grisay, A., E. Gonzalez and **C. Monseur** (2009), *Equivalence of item difficulties across national versions of the PIRLS and PISA reading assessements*, http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02.pdf#page=63 (accessed on 16 July 2019). [3]

Grisay, A. and **C. Monseur** (2007), "Measuring the equivalence of item difficulty in the various versions of an international test", *Studies in Educational Evaluation*, Vol. 33/1, pp. 69-86, http://dx.doi.org/10.1016/j.stueduc.2007.01.006. [2]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [1]

ANNEX A7
Comparing reading, mathematics and science performance across PISA cycles

The methodology underpinning the analysis of trends in performance in international studies of education is complex. In order to ensure the comparability of PISA results across different assessment years, a number of conditions must be met.

In particular, successive assessments of the same subject must include a sufficient number of common assessment items, and these items must retain their measurement properties over time, so that results can be reported on a common scale. The set of items included must adequately cover the different aspects of the framework for each domain.

Furthermore, the sample of students in assessments carried out in different years must be equally representative of the target population; only results from samples that meet the strict standards set by PISA can be compared over time. Even though they participated in successive PISA assessments, some countries and economies cannot compare all of their PISA results over time.

Even when PISA samples accurately reflect the target population (that of 15-year-olds enrolled in grade 7 or above), changes in enrolment rates and demographics can affect the interpretation of trends. For this reason, Chapter 9 in this volume also discusses contextual changes alongside trends in performance, and presents adjusted trends that account for changes in the student population in addition to the basic, non-adjusted performance trends.

Comparisons over time can also be affected by changes in assessment conditions or in the methods used to estimate students' performance on the PISA scale. In particular, from 2015 onward, PISA introduced computer-based testing as the main mode of assessment. It also adopted a more flexible model for scaling response data, and treated items that were left unanswered at the end of test forms as if they were not part of the test, rather than as incorrectly answered. (Such items were considered incorrect in previous cycles for the purpose of estimating students' position on the PISA scale.) Instead of re-estimating past results based on new methods, PISA incorporates the uncertainty associated with these changes when computing the significance of trend estimates (see the section on "link errors" below, and Chapter 2).

Finally, comparisons of assessment results through years that correspond to different assessment frameworks may also reflect the shifting emphasis of the test. For example, differences between PISA 2015 (and earlier) and PISA 2018 results in reading, or between PISA 2012 and PISA 2018 results in science reflect not only whether students have become better at mastering the common assessment items used for linking the assessments (which reflect the earlier assessment framework), they also reflect students' relative performance (compared to other students, in other countries) on aspects of proficiency that are emphasised in the most recent assessment framework.

LINK ERRORS

Link errors are estimates that quantify the uncertainty involved in comparisons that involve different calibrations of the same scale (e.g. the PISA 2009 and the PISA 2018 calibrations of the reading scale). Standard errors for estimates of changes in performance and trends across PISA cycles take this uncertainty into account.

As in past cycles, only the uncertainty around the location of scores from past PISA cycles on the 2018 reporting scale is reflected in the link error. Because this uncertainty about the position in the distribution (a change in the intercept) is cancelled out when looking at location-invariant estimates (such as estimates of the variance, the inter-quartile range, gender gaps, regression coefficients, correlation coefficients, etc.), standard errors for these estimates do not include the linking error.

Link error for scores between two PISA assessments

Link errors for PISA 2018 were estimated based on the comparison of rescaled country/economy means per domain with the corresponding means derived from public use files and produced under the original scaling of each cycle. This approach for estimating the link errors was used for the first time in PISA 2015 (OECD, 2017, p. 237[1]). The number of observations used for the computation of each link error equals the number of countries with results in both cycles. Because of the sparse nature of the data underlying the computation of the link error, a robust estimate of the standard deviation was used, based on the S_n statistic (Rousseeuw and Croux, 1993[2]).

Table I.A7.1 **Link errors for comparisons between PISA 2018 and previous assessments**

Comparison	Reading	Mathematics	Science
PISA 2000 to 2018	4.04		
PISA 2003 to 2018	7.77	2.80	
PISA 2006 to 2018	5.24	3.18	3.47
PISA 2009 to 2018	3.52	3.54	3.59
PISA 2012 to 2018	3.74	3.34	4.01
PISA 2015 to 2018	3.93	2.33	1.51

Note: Comparisons between PISA 2018 scores and previous assessments can only be made to when the subject first became a major domain or later assessment cycles. As a result, comparisons of mathematics and science performance between PISA 2000 and PISA 2018, for example, are not possible.

Source: *PISA 2018 Technical Report* (OECD, forthcoming[3]).

StatLink https://doi.org/10.1787/888934028957

Link error for other types of comparisons of student performance

In PISA, link errors for comparisons across two assessments are considered to be the same across the scale: the link error is the same for a scale score of 400 as for a scale score of 600. However, not all quantities of interest are reported on the PISA scale; and some comparisons involve more than two assessments. How is the proportion of students scoring above a particular cut-off value affected by the link error? How are regression-based trends affected by link errors?

The link error for regression-based trends in performance and for comparisons based on non-linear transformations of scale scores can be estimated by simulation, based on the link error for comparison of scores between two PISA assessments. In particular, Table I.A7.2 (available on line) presents the estimates of the link error for the comparison of the percentage of students performing below Level 2 and at or above Level 5, while Table I.A7.3 presents the magnitude of the link error associated with the estimation of the average three-year trend (see below for a definition of the average three-year-trend).

The estimation of the link errors for the percentage of students performing below Level 2 and at or above Level 5 uses the assumption that the magnitude of the uncertainty associated with the linking of scales follows a normal distribution with a mean of 0 and a standard deviation equal to the scale link error shown in Table I.A7.1. From this distribution, 500 errors are drawn and added to the first plausible value of each country's/economy's 2018 students, to represent the 500 possible scenarios in which the only source of differences with respect to 2018 is the uncertainty in the link.

By computing the estimate of interest (such as the percentage of students in a particular proficiency level) for each of the 500 replicates, it is possible to assess how the scale link error influences this estimate. The standard deviation of the 500 replicate estimates is used as the link error for the change in the percentage of students scoring at a particular proficiency level. Because the influence of the scale link error on this estimate depends on the exact shape and density of the performance distribution around the cut-off points, link errors for comparisons of proficiency levels are different for each country, and within countries, for boys and girls.

The estimation of the link errors for regression-based trends similarly uses the assumption that the uncertainty in the link follows a normal distribution with a mean of 0 and a standard deviation equal to the scale link error shown in Table I.A7.1. However, because the interest here lies in trends over more than two assessment years, the covariance between link errors must be considered in addition to the link errors shown in Table I.A7.1.

To simulate data from multiple PISA assessments, 2 000 observations were drawn from a multivariate normal distribution with all means equal to 0 and whose variance/covariance structure is identified by the link error published in Table I.A7.1, and by those between previous PISA reporting scales, published in Table 12.31 of the *PISA 2012 Technical Report* and in Table 12.8 of the *PISA 2015 Technical Report* (OECD, 2014[4]; OECD, 2017[1]). These draws represent 2 000 possible scenarios in which the real trend is 0, and the estimated trend entirely reflects the uncertainty in the comparability of scores across scales. Link errors for comparisons of the average three-year trend between PISA 2018 and previous assessments depend on the number of cycles involved in the estimation, but are independent of the shape of the performance distribution within each country.

Comparisons of performance: Difference between two assessments and average three-year trend

To evaluate the evolution of performance, analyses report the change in performance between two cycles and the average three-year trend in performance. When more than five data points are available, curvilinear trend trajectories are also estimated.

Comparisons between two assessments (e.g. a country's/economy's change in performance between PISA 2009 and PISA 2018 or the change in performance of a subgroup) are calculated as:

Equation I.A7.1 $$\Delta_{2018-t} = PISA_{2018} - PISA_t$$

where $\Delta_{2018\text{-}t}$ is the difference in performance between PISA 2018 and a previous PISA assessment, $PISA_{2018}$ is the mathematics, reading or science score observed in PISA 2018, and $PISA_t$ is the mathematics, reading or science score observed in a previous assessment. (Comparisons are only possible with the year when the subject first became a major domain or later assessments; as a result, comparisons of mathematics performance between PISA 2018 and PISA 2000 are not possible, nor are comparisons of science performance between PISA 2018 and PISA 2000 or PISA 2003.) The standard error of the change in performance $\sigma(\Delta_{2018\text{-}t})$ is:

Equation I.A7.2 $$\sigma(\Delta_{2018-t}) = \sqrt{\sigma_{2018}^2 + \sigma_t^2 + error_{2018,t}^2}$$

where σ_{2018} is the standard error observed for $PISA_{2018}$, σ_t is the standard error observed for $PISA_t$ and $error^2_{2018,t}$ is the link error for comparisons of science, reading or mathematics performance between the PISA 2018 assessment and a previous (t) assessment. The value for $error^2_{2018,t}$ is shown in Table I.A7.1 for most of the comparisons and Table I.A7.2 for comparisons of proficiency levels.

A second set of analyses reported in this volume relates to the average three-year trend in performance. The average three-year trend is the average rate of change observed through a country's/economy's participation in PISA per three-year period – an interval corresponding to the usual interval between two consecutive PISA assessments. Thus, a positive average three-year trend of x points indicates that the country/economy has improved in performance by x points per three-year period since its earliest comparable PISA results. For countries and economies that have participated only in PISA 2015 and PISA 2018, the average three-year trend is equal to the difference between the two assessments.

The average three-year trend in performance is calculated through a regression of the form

Equation I.A7.3 $$PISA_{i,t} = \beta_0 + \beta_1 time_t + \varepsilon_{i,t}$$

where $PISA_{i,t}$ is country *i*'s location on the science, reading or mathematics scale in year *t* (mean score or percentile of the score distribution), $time_t$ is a variable measuring time in three-year units, and $\varepsilon_{i,t}$ is an error term indicating the sampling and measurement uncertainty around $PISA_{i,t}$. In the estimation, sampling errors and measurement errors are assumed to be independent across time. Under this specification, the estimate for β_1 indicates the average rate of change per three-year period. Just as a link error is added when drawing comparisons between two PISA assessments, the standard errors for β_1 also include a link error:

Equation I.A7.4 $$\sigma(\beta_1) = \sqrt{\sigma_{s,i}{}^2(\beta_1) + \sigma_l{}^2(\beta_1)}$$

where $\sigma_{s,i}$ (β_1) is the sampling and imputation error associated with the estimation of β_1 and σ_l^2 (β_1) is the link error associated with the average three-year trend. It is presented in Table I.A7.3.

The average three-year trend is a more robust measure of a country's/economy's progress in education outcomes as it is based on information available from all assessments. It is thus less sensitive to abnormal measurements that may alter comparisons based on only two assessments. The average three-year trend is calculated as the best-fitting line throughout a country's/economy's participation in PISA. PISA scores are regressed on the year the country participated in PISA (measured in three-year units of time). The average three-year trend also takes into account the fact that, for some countries and economies, the period between PISA assessments is less than three years. This is the case for those countries and economies that participated in PISA 2000 or PISA 2009 as part of PISA+. They conducted the assessment in 2001, 2002 or 2010 instead of 2000 or 2009.[1]

Curvilinear trends are estimated in a similar way, by fitting a quadratic regression function to the PISA results for country *i* across assessments indexed by *t*:

Equation I.A7.5 $$PISA_{i,t} = \beta_2 + \beta_3 year_t + \beta_4 year_t^2 + \varepsilon_{i,t}$$

where $year_t$ is a variable measuring time in years since 2018 and $year_t^2$ is equal to the square of year *t*. Because *year* is scaled such that it is equal to zero in 2018, β_3 indicates the estimated annual rate of change in 2018 and β_4 the acceleration/deceleration of the trend. If β_4 is positive, it indicates that the observed trend is U-shaped, and rates of change in performance observed in years closer to 2018 are higher (more positive) than those observed in earlier years. If β_4 is negative, the observed trend has an inverse-U shape, and rates of change in performance observed in years closer to 2018 are lower (more negative) than those observed in earlier years. Just as a link error is added in the estimation of the standard errors for the average three-year trend, the standard errors for β_3 and β_4 also include a link error (Table I.A7.4). Curvilinear trends are only estimated for countries/economies that can compare their performance across five assessments at least, to avoid over-fitting the data.

ADJUSTED TRENDS

PISA maintains its technical standards over time. Although this means that trends can be calculated over populations defined in a consistent way, the share of the 15-year-old population that this represents, and/or the demographic characteristics of 15-year-old students can also be subject to change, for example because of migration.

Because trend analyses illustrate the pace of progress of successive cohorts of students, in order to draw reliable conclusions from such results, it is important to examine the extent to which they are driven by changes in the coverage rate of the sample and in the demographic characteristics of students included in the sample. Three sets of trend results were therefore developed: unadjusted trends, adjusted trends accounting for changes in enrolment, and adjusted trends accounting for changes in the demographic characteristics of the sample. Adjusted trends represent trends in performance estimated after neutralising the impact of concurrent changes in the demographic characteristics of the sample.

Adjusted trends accounting for changes in enrolment

To neutralise the impact of changes in enrolment rates on trends in median performance and on performance at higher percentiles (or, more precisely, the impact of changes in the coverage rate of the PISA sample with respect to the total population of 15-year-olds; see Coverage Index 3 in Annex A2), the assumption was made that the 15-year-olds not covered by the assessment would all perform below the percentile of interest across all 15-year-olds. With this assumption, the median score across all 15-year-olds (for countries where the coverage rate of the sample is at least 50%) and higher percentiles could be computed without the need to specify the level of performance of the 15-year-olds who were not covered (note that the assumption made is more demanding for the median than for higher percentiles, such as the 75th percentile).

In practice, the estimation of adjusted trends accounting for changes in enrolment first requires that a single case by country/economy be added to the database, representing all 15-year-olds not covered by the PISA sample. The final student weight for this case is computed as the difference between the total population of 15-year-olds (see Table I.A2.2) and the sum of final student weights for the observations included in the sample (the weighted number of participating students). Similarly, each replicate weight for this case is computed as the difference between the total population of 15-year-olds and the sum of the corresponding replicate weights. Any negative weights resulting from this procedure are replaced by 0. A value below any of the plausible values in the PISA sample is entered for the performance variables of this case.

In a second step, the median and upper percentiles of the distribution are computed on the augmented sample. In a few cases where the coverage rate is below 50%, the estimate for the adjusted median is reported as missing.

Adjusted trends accounting for changes in the demographic characteristics of the sample

A re-weighting procedure, analogous to post-stratification, is used to adjust the sample characteristics of past samples to the observed composition of the PISA 2018 sample.

In a first step, the sample included in each assessment cycle is divided into discrete cells, defined by the students' immigrant status (four categories: non-immigrant, first-generation, second-generation, missing), gender (two categories: boy, girl) and relative age (four categories, corresponding to four three-month periods). The few observations included in past PISA datasets with missing gender or age are deleted. This defines, at most, 32 discrete cells for the entire population. However, whenever the number of observations included in one of these 32 cells is less than 10 for a certain country/economy and PISA assessment, the corresponding cell is combined with another, similar cell, according to a sequential algorithm, until all cells reach a minimum sample size of 10.

In a second step, the cells are reweighted so that the sum of final student weights within each cell is constant across assessments, and equal to the sum of final student weights in the PISA 2018 sample. Estimates of the mean and distribution of student performance are then calculated on these reweighted samples, representing the (counterfactual) performance that would have been observed had the samples from previous years had the same composition of the sample in PISA 2018 in terms of the variables used in this re-weighting procedure.

COMPARING THE OECD AVERAGE ACROSS PISA CYCLES

Throughout this report, the OECD average is used as a benchmark. It is calculated as the average across OECD countries, weighting each country equally. Some OECD countries did not participate in certain assessments; other OECD countries do not have comparable results for some assessments; still others did not include certain questions in their questionnaires or changed them substantially from assessment to assessment. In trend tables and figures, the OECD average is reported on consistent sets of OECD countries, and multiple averages may be included. For instance, the "OECD average-23" includes only 23 OECD countries that have non-missing observations for all assessments since PISA 2000; other averages include only OECD countries that have non-missing observations for the years for which this average itself is non-missing. This restriction allows for valid comparisons of the OECD average over time and neutralises the effect of changing OECD membership and participation in PISA on the estimated trends.

Tables available on line

https://doi.org/10.1787/888934028957

- Table I.A7.2. Link errors for comparisons of proficiency levels between PISA 2018 and previous assessments
- Table I.A7.3. Link errors for the linear trend between previous assessments and PISA 2018
- Table I.A7.4. Link errors for the curvilinear trend between previous assessments and PISA 2018

....

Notes

1. Countries and economies that participated in the PISA+ projects administered the same assessments as their PISA 2000 or PISA 2009 counterparts, the only difference being that the assessments were conducted one or two years later. These countries'/economies' data were adjudicated against the same technical and quality standards as their PISA 2000 and PISA 2009 counterparts. Results from the PISA+ projects appeared originally in OECD/UNESCO Institute for Statistics (2003[6]) and Walker (2011[5]), and data from these countries and economies are available as part of the PISA 2000 and PISA 2009 data sets.

References

OECD (2017), *PISA 2015 Technical Report*, OECD Publishing, Paris, http://www.oecd.org/pisa/data/2015-technical-report/ (accessed on 31 July 2017). [1]

OECD (2014), *PISA 2012 Technical Report*, OECD Publishing, Paris, http://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf (accessed on 18 September 2019). [4]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [3]

OECD/UNESCO Institute for Statistics (2003), *Literacy Skills for the World of Tomorrow: Further Results from PISA 2000*, OECD Publishing, http://dx.doi.org/10.1787/9789264102873-en. [6]

Rousseeuw, P. and C. Croux (1993), "Alternatives to the Median Absolute Deviation", *Journal of the American Statistical Association*, Vol. 88/424, pp. 1273–83, http://dx.doi.org/10.1080/01621459.1993.10476408. [2]

Walker, M. (2011), *PISA 2009 Plus Results: Performance of 15-year-olds in reading, mathematics and science for 10 additional participants*, ACER Press. [5]

ANNEX A8
How much effort did students invest in the PISA test?

Performance on school tests is the result of the interplay amongst what students know and can do, how quickly they process information, and how motivated they are for the test. To ensure that students who sit the PISA test engage with the assessment conscientiously and sustain their efforts throughout the test, schools and students that are selected to participate in PISA are often reminded of the importance of the study for their country. For example, at the beginning of the test session, the test administrator reads a script that includes the following sentence:

> *"This is an important study because it will tell us about what you have been learning and what school is like for you. Because your answers will help influence future educational policies in <country and/or education system>, we ask you to do the very best you can."*

However, viewed in terms of the individual student who takes the test, PISA can be described as a low-stakes assessment: students can refuse to participate in the test without suffering negative consequences, and do not receive any feedback on their individual performance. If students perceive an absence of personal consequences associated with test performance, there is a risk that they might not invest adequate effort (Wise and DeMars, 2010[1]).

Several studies in the United States have found that student performance on assessments, such as the United States national assessment of educational progress (NAEP), depends on the conditions of administration. In particular, students performed less well in regular low-stakes conditions compared to experimental conditions in which students received financial rewards tied to their performance or were told that their results would count towards their grades (Wise and DeMars, 2005[2]). In contrast, a study in Germany found no difference in effort or performance measures between students who sat a PISA-based mathematics test under the standard PISA test-administration conditions, and students who sat the test in alternative conditions that increased students' stakes for performing well (Baumert and Demmrich, 2001[3]). In the latter study the experimental conditions included promising feedback about one's performance, providing monetary incentives contingent on performance, and letting students know that the test would count towards their grades. The difference in results may suggest that the motivation of students to expend effort on a low-stakes test such as PISA may differ significantly across countries. Indeed, the only existing comparative study on the effect of incentives on test performance found that offering students monetary incentives to expend effort on a test such as PISA – something that is not possible within the regular PISA procedures – led to improved performance amongst students in the United States, while students in Shanghai (China) performed equally well with or without such incentives (Gneezy et al., 2017[4]).

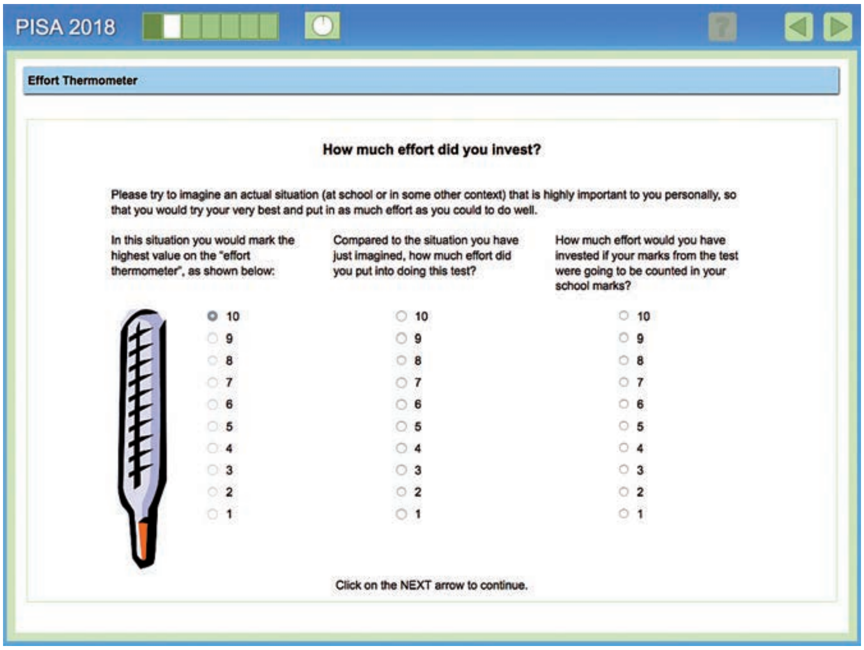
These studies suggest that differences in countries' and economies' mean scores in PISA may reflect differences not only in what students know and can do, but also in their motivation to do their best. Put differently, PISA does not measure students' maximum potential, but what students actually do, in situations where their individual performance is monitored only as part of their group's performance.

A number of indicators have been developed to assess differences between individuals or between groups (e.g. across countries and economies) in students' motivation in low-stakes tests.

Several scholars have used student self-report measures, collected shortly after the test (Wise and DeMars, 2005[2]; Eklöf, 2007[5]). Typically, students are asked about the effort they invested in the test, and the effort they would have expended in a hypothetical situation, e.g. if the test results counted towards their grades. PISA 2018 also included such questions at the end of both the paper- and computer-based test forms (see Figure I.A8.1).

However, there are several disadvantages of self-report measures. In particular, it is unclear whether students – especially those who may not have taken the test seriously – respond truthfully when asked how hard they tried on a test they have just taken; and it is unclear to what extent answers provided on subjective response scales can be compared across students, let alone across countries. The comparison between the "actual" and the "hypothetical" effort is also problematic. In the German study discussed above, regardless of the conditions under which they took the test, students said that they would have invested more effort if any of the other three conditions applied; the average difference was particularly marked amongst boys (Baumert and Demmrich, 2001[3]). One explanation for this finding is that students are under-reporting their true effort, relative to their hypothetical effort, to attribute their wrong answers in the actual test they sat to lack of effort, rather than lack of ability.

Figure I.A8.1 **The effort thermometer in PISA 2018**



In response to these criticisms, researchers have developed new ways to examine test-taking effort, based on the observation of students' behaviour during the test. Wise and Kong (2005[6]) propose a measure based on response time per item in computer-delivered tests. Their measure, labelled "response-time effort", is simply the proportion of items, out of the total number of items in a test, on which respondents spent more than a threshold time T (e.g five seconds, for items based on short texts). Borgonovi and Biecek (2016[7]) developed a country-level measure of "academic endurance" based on comparisons of performance in the first quarter of the PISA 2012 test and in the third quarter of the PISA 2012 test (the rotating booklets design used in PISA 2012 ensured that test content was perfectly balanced across the first and third quarters). The reasoning behind this measure is that while effort can vary during the test, what students know and can do remains constant; any difference in performance is therefore due to differences in the amount of effort invested.[1] Measures of "straightlining", i.e. the tendency to use an identical response category for all items in a set (Herzog and Bachman, 1981[8]), may also indicate low test-taking effort.

Building on these measures, this annex presents country-level indicators of student effort and time management in PISA 2018, and compares them, where possible, to the corresponding PISA 2015 measures. The intention is not to suggest adjustments to PISA mean scores or performance distributions, but to provide richer context for interpreting cross-country differences and trends.

AVERAGE STUDENT EFFORT AND MOTIVATION

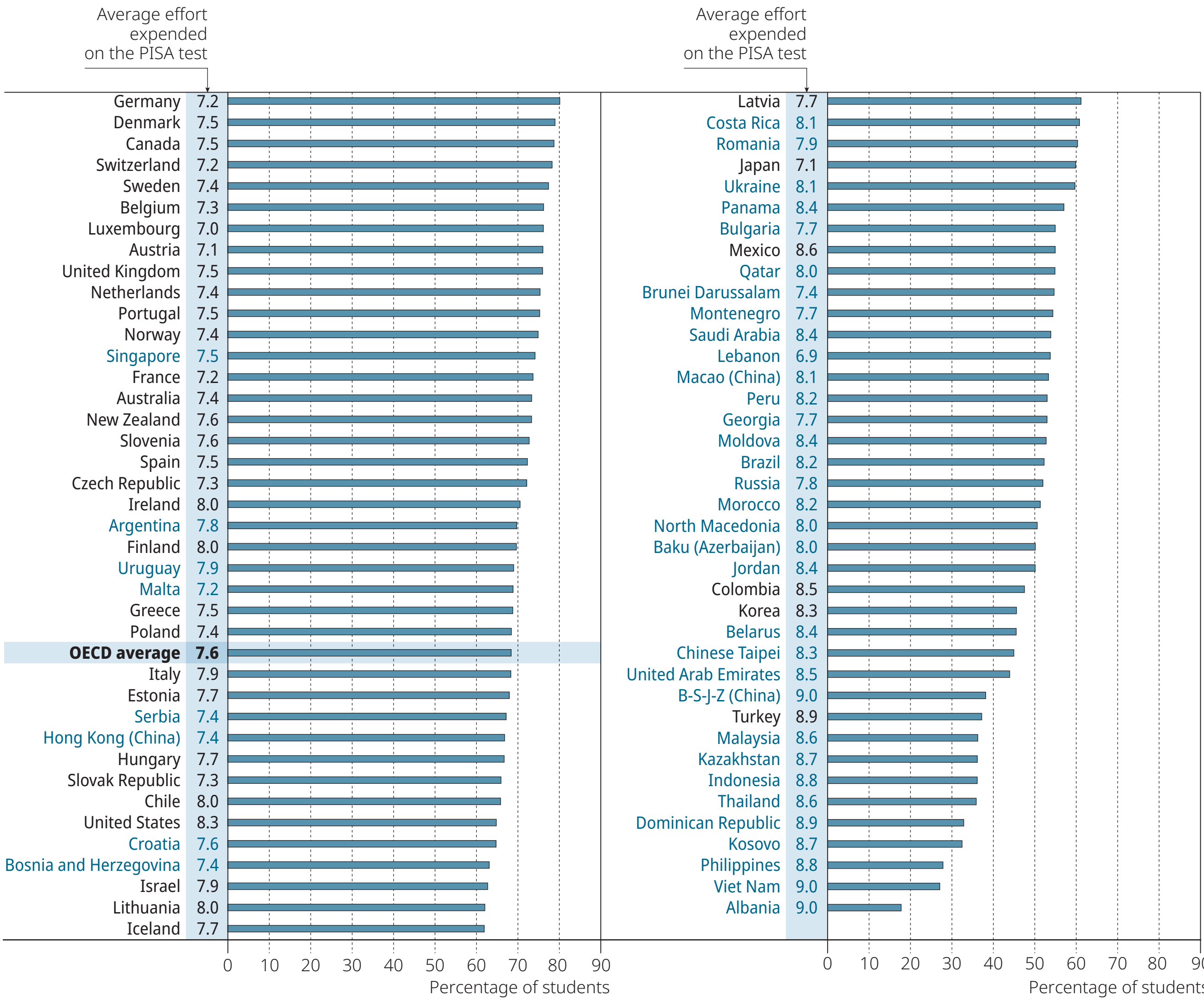
Figure I.A8.2 presents the results from students' self-reports of effort; Figure I.A8.3 presents, for countries using the computer-based test, the result of Wise and Kong's (2005[6]) measure of effort based on item-response time.

A majority of students across OECD countries (68%) reported expending less effort on the PISA test than they would have done in a test that counted towards their grades. On the 1-to-10 scale illustrated in Figure I.A8.1, students reported an effort of "8" for the PISA test they just had completed, on average. They reported that they would have described their effort as "9" had the test counted towards their marks. Students in Albania, Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]") and Viet Nam rated their effort highest, on average across all participating countries/economies, with an average rating of "9". Only 17% of students in Albania and 27% of students in Viet Nam reported that they would have invested more effort had the test counted towards their marks.

At the other extreme, more than three out of four PISA students in Germany, Denmark, Canada, Switzerland, Sweden, Belgium, Luxembourg, Austria, the United Kingdom, the Netherlands and Portugal (in descending order of that share), and 68% on average across OECD countries, reported that they would have invested more effort if their performance on the test had counted towards their marks (Table I.A8.1). In most countries, as well as on average, boys reported investing less effort in the PISA test than girls reported. But the effort that boys reported they would have invested in the test had it counted towards their marks was also less than the effort that girls reported under the same hypothetical conditions. When the difference between the reports is considered, a larger share of girls reported that they would have worked harder on the test if it had counted towards their marks (Table I.A8.2).

Figure I.A8.2 **Self-reported effort in PISA 2018**

Percentage of students who reported expending less effort on the PISA test than if the test counted towards their marks



Note: The number next to the name of the country/economy indicates the average effort expended on the PISA test, reported on a 0-to-10 scale by students.

Countries and economies are ranked by descending order of percentage of students who reported expending less effort on the PISA test than if the test counted towards their marks.

Source: OECD, PISA 2018 Database, Table I.A8.1.

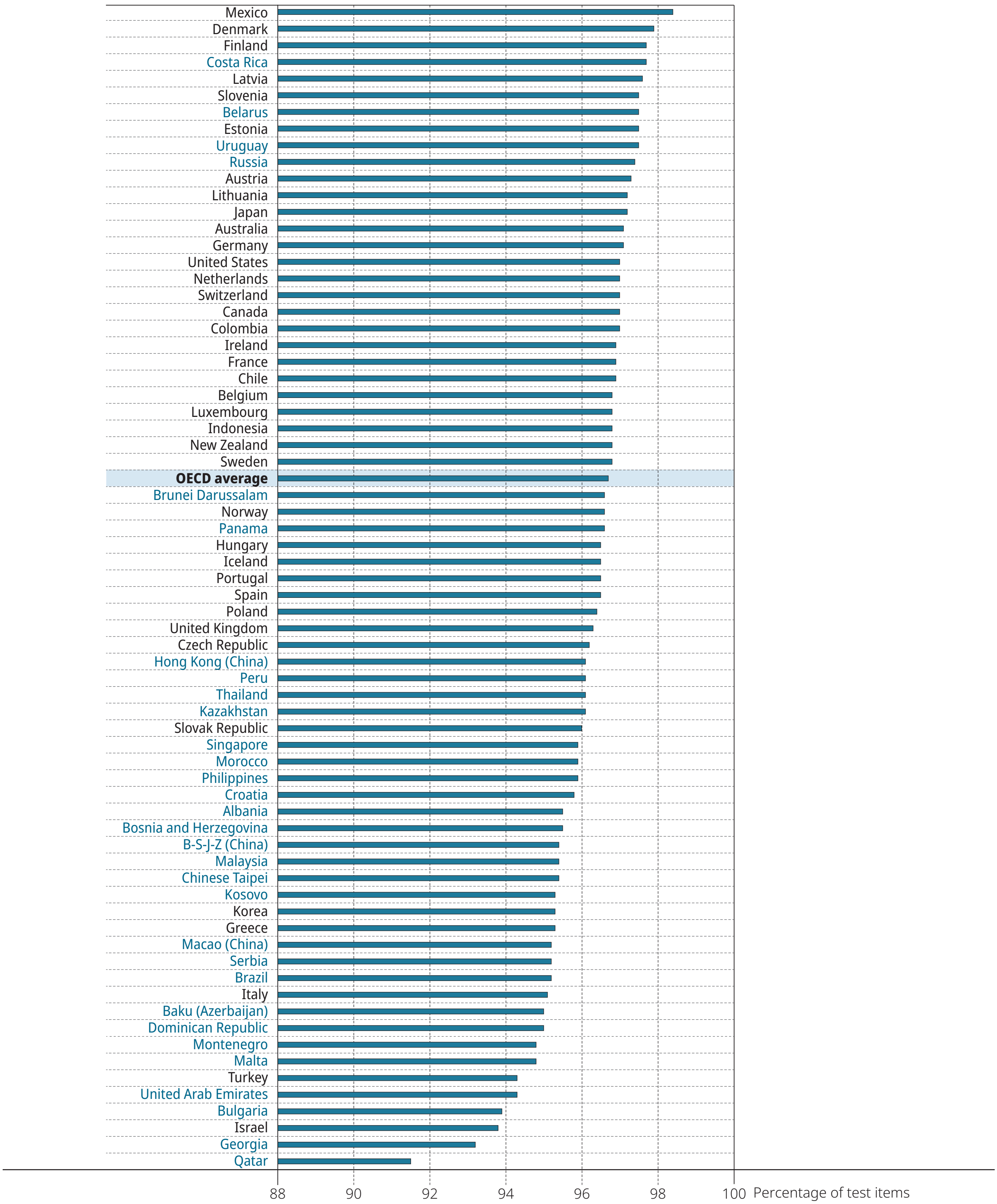
StatLink https://doi.org/10.1787/888934028976

Response-time effort, on the other hand, appears unrelated to country-level ratings of self-reported effort (this measure is available only for countries that delivered PISA on computer).[2] In fact, most countries/economies show considerable response-time effort. In order to estimate response-time effort, a conservative threshold (i.e. a minimum of five seconds per item) was used to define “solution behaviour” on mathematics and science items; reading and global competence items were excluded from the items considered to ensure comparability across countries.[3]

The largest share of students who exhibited genuine response behaviour – i.e. spent at least five seconds on any mathematics or science item that was presented to them – were found in Denmark (one of the countries with the largest share of students who reported they would have worked harder on the test if it had counted towards their marks), Finland and Mexico; but many other countries and economies share similar levels of response-time effort as these three. Only Qatar (response-time effort equal to 91.5%) has a large share of student responses to items (8.5%) that may correspond to “rapid skipping” or “rapid guessing” behaviours (i.e. students spent less than five seconds trying to solve the item; rapidly skipped items at the end of each session were considered as non-reached items and did not count towards the measure of response-time effort; Table I.A8.7). The same pattern of low response-time effort in Qatar was already observed in 2015 data (Table I.A8.9).[4]

Figure I.A8.3 **Response-time effort in PISA 2018**

Average percentage of test items on which students spent more than 5 seconds (excluding reading and global competence items)



Countries and economies are ranked in descending order of the percentage of items on which students spent more than 5 seconds.

Source: OECD, PISA 2018 Database, Table I.A8.7.

StatLink https://doi.org/10.1787/888934028995

One possibility is that the differences between self-report and response-time measures of effort arise because response-time effort is not sensitive to all types of disengaged response behaviours. Not all students who expend little effort skip questions or give rapid-guess answers; some may get off task, or read through the material without focus, and end up guessing or skipping responses only after expending considerable time. Another possibility is that self-report measures of effort do not reflect the real effort that students put into the test (see above).

The reading-fluency section of the PISA 2018 test offers an opportunity to examine straightlining behaviour in the test. Students were given a series of 21 or 22 items, in rapid sequence, with identical response formats ("yes" or "no"); meaningless sentences (such as "The window sang the song loudly"), calling for a "no" answer, were interspersed amongst sentences that had meaning (such as "The red car has a flat tyre"), calling for a "yes" answer. It is possible that some students did not read the instructions carefully, or that they genuinely considered that the meaningless sentences (which had no grammatical or syntactical flaws) had meaning. However, this response pattern (a series of 21 or 22 "yes" answers) or its opposite (a series of 21 or 22 "no" answers) is unexpected amongst students who demonstrated medium or high reading competence in the main part of the reading test.

Table I.A8.21 shows that, indeed, only 1.5% of all students, on average across OECD countries, exhibited such patterned responses in reading-fluency tasks. That proportion is even smaller (only 0.5%) amongst high-performing students, defined here as those students who attained high scores on the first segment of the reading test, after completing reading-fluency tasks.[5] However, the proportion of high-performing students who exhibited "straightlining" behaviour on the reading-fluency test is close to 6% in Kazakhstan, close to 5% in the Dominican Republic, and exceeds 2% in Albania, Indonesia, Korea, Peru, Spain, Thailand and Turkey (Table I.A8.21).[6] It is possible that the unusual response format of reading-fluency tasks triggered, in some limited cases, disengaged response behaviour, and that these same students did their best in the later parts of the test. It is also possible, however, that these students did not do their best throughout the PISA test, and not only in this initial, three-minute section of the reading test.

TEST FATIGUE AND THE ABILITY TO SUSTAIN MOTIVATION

For countries that delivered the test on computer, Figure I.A8.4 presents a measure of test endurance based on Borgonovi and Biecek (2016[7]). This measure compares performance in mathematics and science (domains where the assignment of tasks to students was non-adaptive) between the first and second test session (each test session corresponds to one hour). In PISA 2018, there were no students who took mathematics and science tasks in both test sessions; the comparison is therefore between equivalent groups of students, as defined by the random assignment of students to test forms. The rotation of items over test forms further ensures a balanced design.

Amongst countries that delivered the PISA test on computer, only negative or non-significant differences in performance were observed between the second and first hour of testing. This is expected, as these differences mostly reflect fatigue, and they support the interpretation of these results as indicators of "endurance". While the differences tended to be small, in general, seven countries/economies showed a decline of more than three percentage points in the percentage of correct answers between the first and second hour of testing (in ascending order, from the smallest to the largest decline): Chile, Serbia, Baku (Azerbaijan), Colombia, Australia, Norway and Uruguay (Figure I.A8.4 and Table I.A8.3).

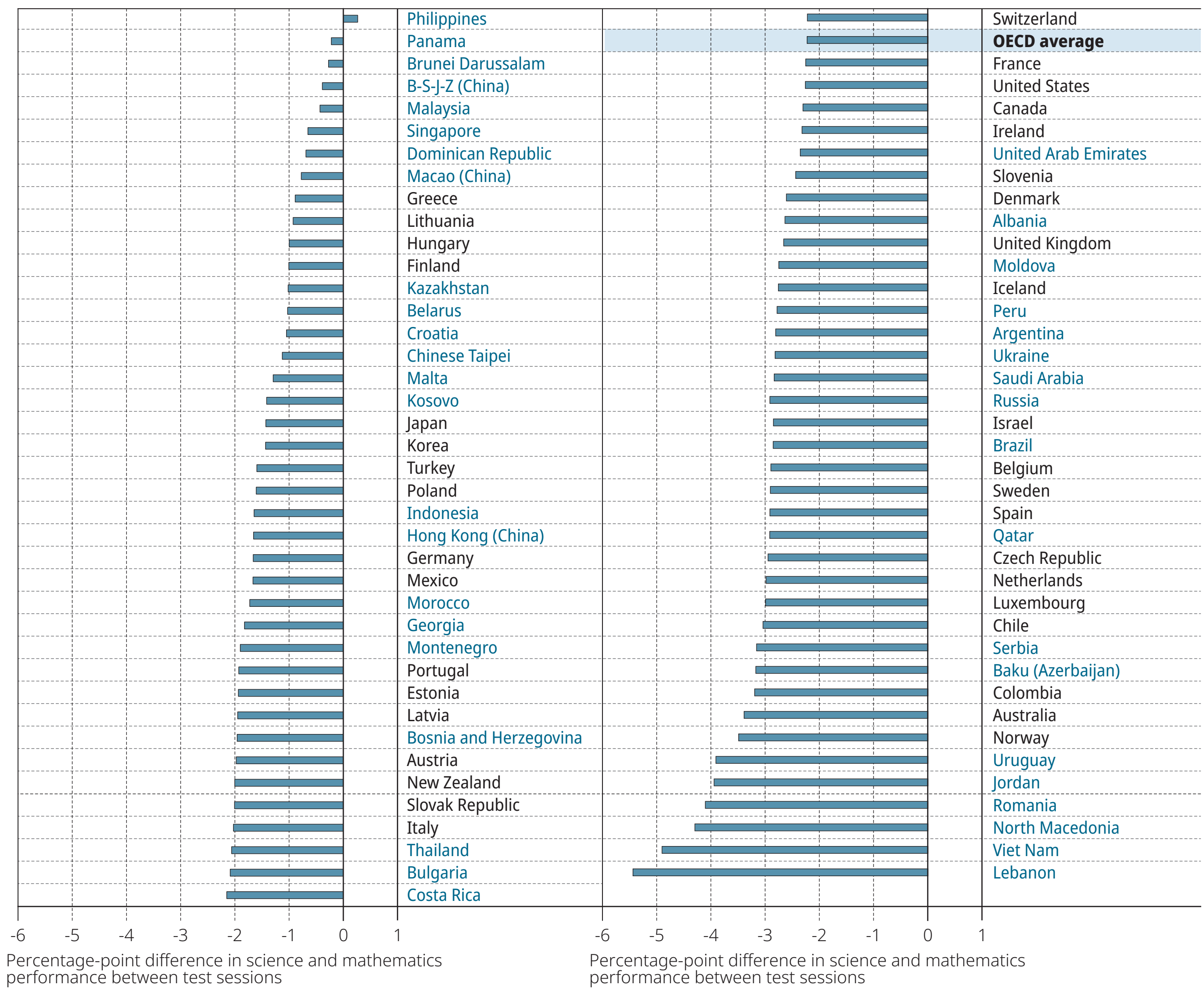
There was hardly a correlation between overall performance and test endurance.[7] Some countries with similar performance in science and mathematics showed marked differences in test endurance. For example, amongst high-performing countries, the Netherlands showed a relatively marked drop in performance between the first and second session, while there was no significant decline in performance in B-S-J-Z (China), Singapore, Macao (China) and Finland (in descending order of the overall percentage of correct responses in mathematics and science). Test endurance was also only weakly related to the share of students in the country who reported expending less effort in the PISA test than if the test counted towards their school grades.[8]

Countries that delivered the PISA test on paper showed, on average, larger differences in percent-correct responses between the first half and the second half of the test. This reflects the different design of the test sessions (see Annex A5). In these countries, students could continue to work on the first half of the test during the second hour of testing, as all domains of testing were bundled in a single booklet.

Academic endurance can be computed in much the same way with PISA 2015 data. In order to compare results with PISA 2018, TableI.A8.5 uses only student performance in mathematics and science. Even so, results should not be compared directly with results in Table I.A8.3, because the test content in science and the distribution of science and mathematics questions across test forms differ between 2015 and 2018 (science was the major domain in 2015, and was always assessed over a full hour). Nevertheless, the PISA 2015 measure of academic endurance correlates strongly at the country level with the PISA 2018 measure of academic endurance (the linear correlation coefficient is $r = 0.65$ across the 53 countries that delivered the test on computer, and that had already delivered the test on computer in 2015) (Tables I.A8.3 and I.A8.5). In general, countries/economies where students showed above-average endurance in 2018 (such as Finland, Macao [China], Singapore and Chinese Taipei) had already demonstrated above-average endurance in 2015, and countries with below-average endurance (such as Australia, the Netherlands, Norway and Uruguay) tended to show below-average endurance in 2015 as well.

Figure I.A8.4 **Academic endurance**

Difference in science and mathematics performance (percent-correct) between test sessions



Countries and economies are ranked in descending order of the difference in science and mathematics performance between test sessions.

Source: OECD, PISA 2018 Database, Table I.A8.3.

StatLink https://doi.org/10.1787/888934029014

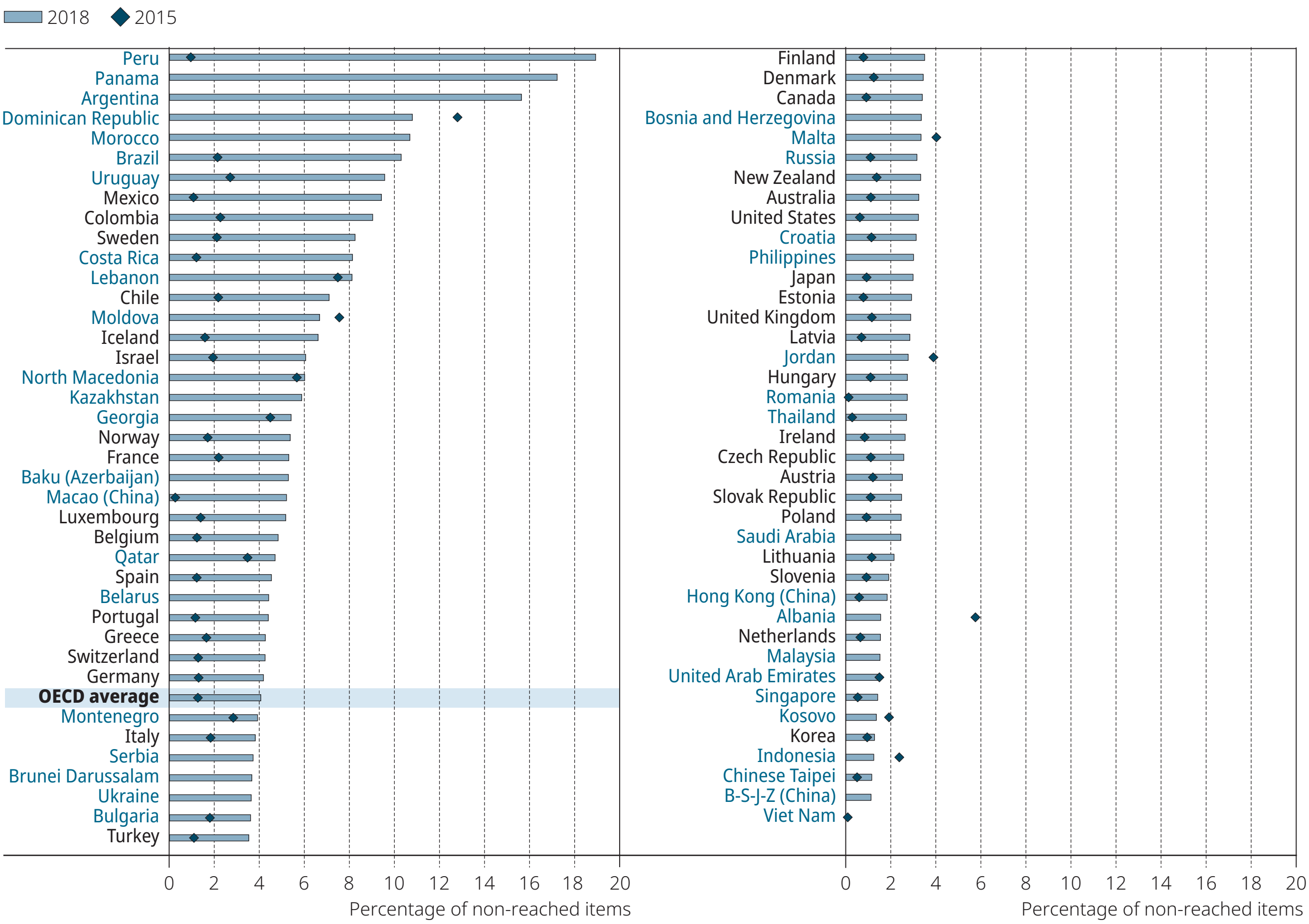
TIME MANAGEMENT AND SPEED OF INFORMATION PROCESSING

Non-reached items at the end of each of the two one-hour test sessions in the computer-based assessment (and at the end of the test booklet, in the paper-based assessment) are defined for each test-taker as omitted responses that are not followed by a valid (correct or incorrect) response before the end of the session/booklet (OECD, forthcoming[9]).

Figure I.A8.5 shows the average percentage of non-reached items in mathematics and science (reading was not analysed due to the adaptive design, which makes the percentage of non-reached items not comparable across students and countries). On average across OECD countries, 4% of items were not reached by the end of the test session: 5% amongst students who were given science or mathematics tests during the first hour, and 3% amongst students who were given science or mathematics tests during the second hour. This difference between the first and second hour, which can be found in most countries that delivered the test on computer, suggests that students may have become more familiar with the test platform, timing and response formats during the test. However, the percentage of non-reached items is above 15% in Peru, Panama and Argentina (in descending order of that percentage; the latter country delivered the test on paper), and it is between 10% and 11% in Brazil, the Dominican Republic and Morocco. The proportion of non-reached items is smallest in Viet Nam (0.1%), followed by B-S-J-Z (China), Korea and Chinese Taipei (between 1.1% and 1.3%) (Figure I.A8.5 and Table I.A8.11).

Figure I.A8.5 **Non-reached items**

Percentage of non-reached items in PISA 2015 and PISA 2018



Notes: Albania, Georgia, Indonesia, Jordan, Kazakhstan, Kosovo and Malta used pen-and-paper tests in 2015 and computer-based tests in 2018. Jordan, Lebanon, the Republic of Moldova, the Republic of North Macedonia, Romania and Viet Nam used pen-and-paper tests in both years. Argentina, Saudi Arabia and Ukraine used pen-and-paper tests in 2018. Results for all other countries are based on computer-based tests.

Countries and economies are ranked in descending order of the percentage of non-reached items in PISA 2018.

Source: OECD, PISA 2018 Database, Tables I.A8.11 and I.A8.13.

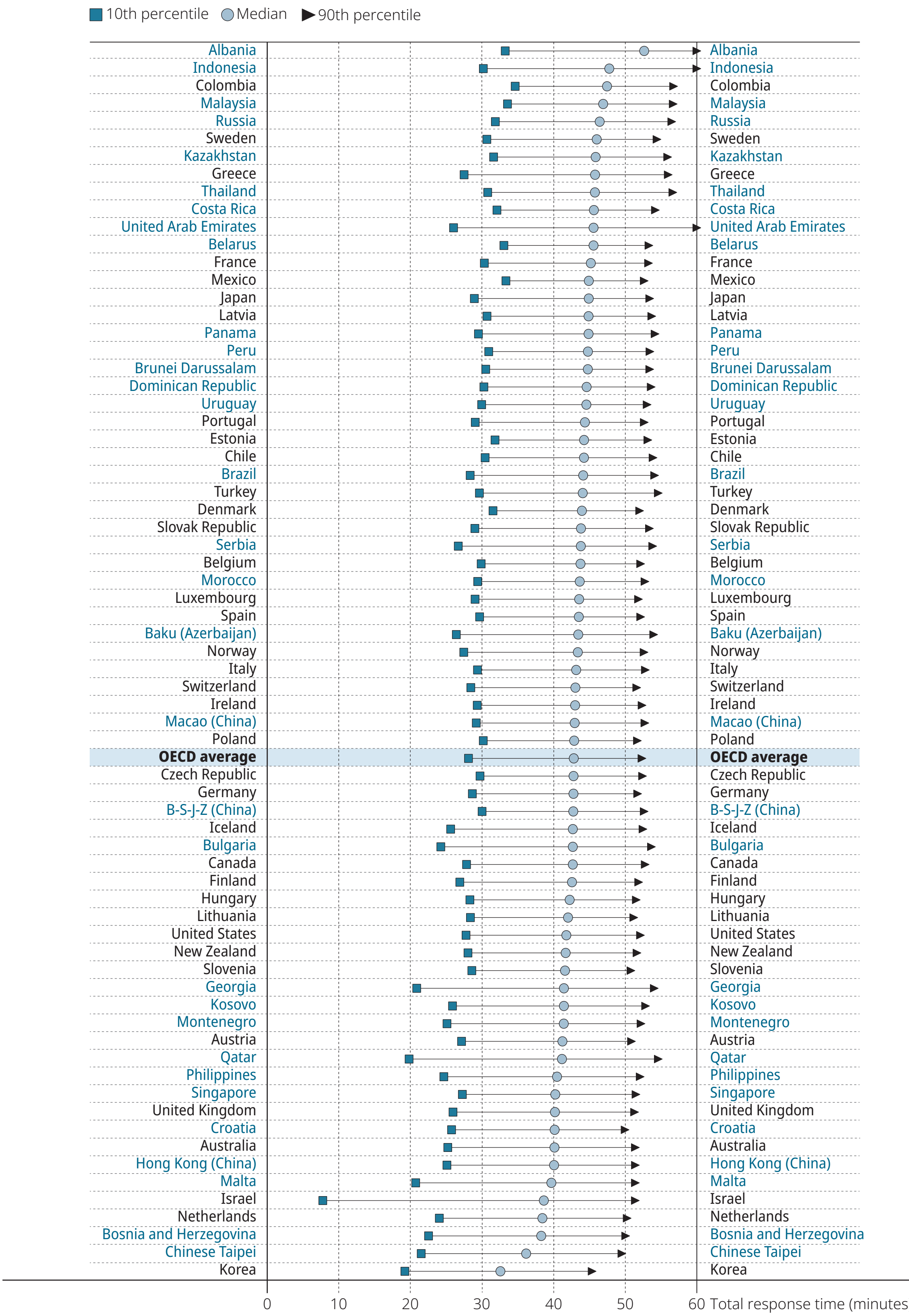
StatLink https://doi.org/10.1787/888934029033

Between 2015 and 2018, the proportion of non-reached items increased in most countries. In many Latin American countries (Brazil, Colombia, Costa Rica, Mexico, Peru and Uruguay), as well as in Sweden, it increased from less than 3% in 2015 to more than 8% in 2018. The most significant exception to this increase is the Dominican Republic, where non-reached items decreased from 13% to 11%. Non-reached items also decreased in most countries that transitioned to computer-based testing in 2018 (Figure I.A8.5; Tables I.A8.11 and I.A8.13). The rotation of the major domain, and other changes affecting the length of the test, may have contributed to the increase in non-reached items in countries that delivered the test on computer. As in 2015, non-reached items were considered as "not administered" for the purpose of estimating students' performance on the PISA scale, and the increase or decrease in non-reached items therefore cannot explain performance changes between 2015 and 2018 (though both changes may be related to the same cause, such as weaker motivation amongst students to try their best).

Figure I.A8.6 presents, for countries that delivered the PISA test on computer, the amount of time students spent on the reading, mathematics and science test. Students were given a maximum of one hour to complete the mathematics and/or science section of their PISA test (the other hour was used for assessing reading). On average across OECD countries, 50% completed the first test section (either the reading section, or the mathematics and/or science section) within less than 43 minutes (median total time); 10% of students took less than 28 minutes to finish the test (10th percentile of total time), and 90% of students completed the test within 52 minutes. Students tended to be faster during the second hour, probably because they became more familiar with the test platform and the different response formats. The median total time was only 39 minutes in the second hour.

Figure I.A8.6 **Overall response time**

Distribution of total response time during the first hour of testing



Countries and economies are ranked in descending order of median total time during the first hour of testing.

Source: OECD, PISA 2018 Database, Table I.A8.15.

StatLink https://doi.org/10.1787/888934029052

Compared to the OECD average, students were considerably faster at completing the test in Korea (median total time: 33 minutes in the first hour, 30 minutes in the second hour). They were considerably slower in Albania (53 minutes in the first hour, 45 minutes in the second hour) and in Malaysia (47 minutes and 46 minutes). In all countries and economies, the vast majority of students completed the test within the time limit (Table I.A8.15).

These patterns of variation across countries in time spent on the test were similar to those observed in 2015 (Table I.A8.17). Across countries/economies with available data, the median total time in the first hour correlates at $r = 0.86$ at the country level. The median test completion time in 2015 was slightly less than in 2018, on average across OECD countries (40 minutes, instead of 43 minutes), suggesting that the PISA 2015 reading and science tests could be completed in less time compared to the PISA 2018 tests (the same mathematics tests were used in 2018 as in 2015). This also aligns with findings that the number of non-reached items increased since 2015.

Tables available on line

https://doi.org/10.1787/888934029071

- Table I.A8.1 Effort invested in the PISA test
- Table I.A8.2 Effort invested in the PISA test, by gender
- Table I.A8.3 Endurance in the PISA test
- Table I.A8.4 Endurance in the PISA test, by gender
- Table I.A8.5 Endurance in the PISA 2015 test
- Table I.A8.6 Endurance in the PISA 2015 test, by gender
- Table I.A8.7 Response-time effort in the PISA test
- Table I.A8.8 Response-time effort in the PISA test, by gender
- Table I.A8.9 Response-time effort in the PISA 2015 test
- Table I.A8.10 Response-time effort in the PISA 2015 test, by gender
- Table I.A8.11 Non-reached items in the PISA test
- Table I.A8.12 Non-reached items in the PISA test, by gender
- Table I.A8.13 Non-reached items in the PISA 2015 test
- Table I.A8.14 Non-reached items in the PISA 2015 test, by gender
- Table I.A8.15 Response time in the PISA test
- Table I.A8.16 Response time in the PISA test, by gender
- Table I.A8.17 Response time in the PISA 2015 test
- Table I.A8.18 Response time in the PISA 2015 test, by gender
- Table I.A8.19 Response time in the PISA reading-fluency test
- Table I.A8.20 Response time in the PISA reading-fluency test, by gender
- Table I.A8.21 Response accuracy in the PISA reading-fluency test, by reading performance
- Table I.A8.22 Response accuracy in the PISA reading-fluency test, by gender

Notes

1. Speed of information processing, and time management more generally, may also influence performance differences between test sections. To limit the influence of this possible confounder, Borgonovi and Biecek (2016[7]) do not use the last quarter of the test, but the third (second-to-last) quarter. In the computer-based PISA 2015 and PISA 2018 assessments, the test is divided in two halves, each conducted in an hour-long session. Under this design, students' time management and speed of information processing can be expected to have the same impact on both halves.

2. The linear correlation coefficient between average response-time effort and self-reported effort in the PISA test is weak (r = -0.20, N = 70). The linear correlation between average response-time effort and the share of students reporting that they invested less effort in the PISA test than if their scores were going to be counted in their school marks is r = 0.38 (N = 70), meaning that in countries with greater response-time effort, more students tended to report that they would have worked harder if the test had had higher stakes for them (Tables I.A8.1 and I.A8.7).

3. In particular, reading items were excluded because their assignment to students was, in part, a function of students' behaviour in prior sections of the test. As a result, each item was assigned to a different proportion of students across countries, limiting comparability of test-wide timing measures. Global competence items were excluded due to the large number of countries that did not participate in the assessment of global competence.

4. More generally, the linear correlation coefficient between response-time effort in 2015 and response-time effort in 2018, at the country level and across the 53 countries/economies that delivered both PISA tests on computers is r = 0.64.

5. High-performing students correctly answered a sufficient number of automatically scored tasks in the core section of the reading test to be assigned, with 90% probability, to a "high" stage-1 testlet in the following section of the adaptive reading test. The same cut-off values (specific to each core testlet) were used across all countries to identify high-performing students. This information is available in variable RCORE_PERF in the PISA 2018 cognitive response database.

6. In all countries and economies, the proportion of correct responses to reading-fluency tasks was positively related to the proportion of correct responses in the core stage of the reading assessment.

7. The linear correlation coefficient between average academic endurance and mean performance in the PISA test is only r = 0.10 in reading, r = 0.13 in mathematics and r = 0.12 in science (N = 78) across all countries/economies. When countries that delivered the PISA test on paper are excluded, correlations are r = -0.08 in reading, r = -0.03 in mathematics and r = -0.03 in science (N = 70) (Tables I.B1.4, I.B1.5, I.B1.6 and I.A8.3).

8. When countries that delivered the PISA test on paper are excluded, the linear correlation coefficient between average academic endurance and the percentage of students who reported that they invested less effort in the PISA test than if their scores were going to be counted in their school marks is r = -0.37 (N = 70), meaning that in countries with better student endurance, a smaller proportion of students indicated that they would have worked harder if the stakes had been higher (Tables I.A8.1 and I.A8.3).

References

Baumert, J. and **A. Demmrich** (2001), "Test motivation in the assessment of student skills: The effects of incentives on motivation and performance", *European Journal of Psychology of Education*, Vol. 16/3, pp. 441-462, http://dx.doi.org/10.1007/bf03173192. [3]

Borgonovi, F. and **P. Biecek** (2016), "An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test", *Learning and Individual Differences*, Vol. 49, pp. 128-137, http://dx.doi.org/10.1016/j.lindif.2016.06.001. [7]

Eklöf, H. (2007), "Test-Taking Motivation and Mathematics Performance in TIMSS 2003", *International Journal of Testing*, Vol. 7/3, pp. 311-326, http://dx.doi.org/10.1080/15305050701438074. [5]

Gneezy, U. et al. (2017), *Measuring Success in Education: The Role of Effort on the Test Itself*, National Bureau of Economic Research , Cambridge, MA, http://dx.doi.org/10.3386/w24004. [4]

Herzog, A. and **J. Bachman** (1981), "Effects of questionnaire length on response quality", *Public Opinion Quarterly*, Vol. 45, pp. 549–559. [8]

OECD (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [9]

Wise, S. and **C. DeMars** (2010), "Examinee Noneffort and the Validity of Program Assessment Results", *Educational Assessment*, Vol. 15/1, pp. 27-41, http://dx.doi.org/10.1080/10627191003673216. [1]

Wise, S. and **C. DeMars** (2005), "Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions", *Educational Assessment*, Vol. 10/1, pp. 1-17, http://dx.doi.org/10.1207/s15326977ea1001_1. [2]

Wise, S. and **X. Kong** (2005), "Response Time Effort: A New Measure of Examinee Motivation in Computer-Based Tests", *Applied Measurement in Education*, Vol. 18/2, pp. 163-183, http://dx.doi.org/10.1207/s15324818ame1802_2. [6]

ANNEX A9
A note about Spain in PISA 2018

Spain's data met PISA 2018 Technical Standards. However, some data show implausible response behaviour amongst students. Consequently, at the time of publication of this report, the OECD is unable to assure that international, subnational and trend comparisons of Spain's results lead to valid conclusions about students' reading proficiency and, more generally, about the education system in Spain. PISA 2018 reading results for Spain are therefore not published in this report and are not included in OECD average results.

The most visible anomalies in students' response behaviour in Spain can be summarised as follows:

- A large number of Spanish students responded to a section of the reading test (the reading-fluency section) in a manner that was obviously not representative of their true reading competency. The assessment is computer based and students' actions are recorded and tracked. A significant proportion of students (including students who scored at high levels in the remaining sections of the PISA test) rushed through the reading-fluency section, spending less than 25 seconds in total over more than 20 test items.
- Many of these students gave patterned responses (all yes or all no, etc.).
- Rapid and patterned responses were not uniformly present in the Spanish sample, but observed predominantly in a small number of schools in some areas of Spain.

The extent and concentration of rapid and patterned responses are unique to Spain, and affect reading performance. Results in the mathematics and science domains appear less affected by anomalous response behaviour, and are therefore published in this report.

The extent and the causes of the anomalies observed are being further investigated to determine if other parts of the reading, science or mathematics test were also affected. The online version of this Annex, available at www.oecd.org/pisa, provides the most current overview of the results of this investigation.

ANNEX B

PISA 2018 Data

All tables in Annex B are available on line

Annex B1: Results for countries and economies
https://doi.org/10.1787/888934029090

Annex B2: Results for regions within countries
https://doi.org/10.1787/888934029109

Annex B3: PISA 2018 system-level indicators
https://doi.org/10.1787/888934029128

ANNEX B1
Results for countries and economies

Table I.B1.1 [1/2] **Percentage of students at each proficiency level in reading**

	All students																	
	Below Level 1c (less than 189.33 score points)		Level 1c (from 189.33 to less than 262.04 score points)		Level 1b (from 262.04 to less than 334.75 score points)		Level 1a (from 334.75 to less than 407.47 score points)		Level 2 (from 407.47 to less than 480.18 score points)		Level 3 (from 480.18 to less than 552.89 score points)		Level 4 (from 552.89 to less than 625.61 score points)		Level 5 (from 625.61 to less than 698.32 score points)		Level 6 (above 698.32 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD																		
Australia	0.1	(0.1)	1.4	(0.2)	5.6	(0.3)	12.5	(0.4)	21.1	(0.5)	25.4	(0.5)	20.9	(0.5)	10.3	(0.4)	2.7	(0.2)
Austria	0.0	(0.0)	0.9	(0.2)	6.4	(0.6)	16.3	(0.8)	23.5	(0.8)	26.2	(0.9)	19.3	(0.8)	6.7	(0.5)	0.7	(0.1)
Belgium	0.1	(0.1)	1.2	(0.2)	6.0	(0.4)	14.0	(0.6)	22.4	(0.7)	26.5	(0.7)	20.4	(0.7)	8.3	(0.5)	1.3	(0.2)
Canada	0.0	(0.0)	0.7	(0.1)	3.1	(0.2)	10.0	(0.4)	20.1	(0.6)	27.2	(0.5)	24.0	(0.5)	12.2	(0.5)	2.8	(0.2)
Chile	0.1	(0.1)	1.7	(0.2)	8.9	(0.6)	21.0	(0.9)	29.5	(0.9)	24.4	(0.9)	11.8	(0.6)	2.4	(0.3)	0.2	(0.1)
Colombia	0.2	(0.1)	3.6	(0.4)	15.8	(0.9)	30.3	(1.0)	27.7	(1.0)	15.8	(0.9)	5.7	(0.5)	0.9	(0.2)	0.0	(0.0)
Czech Republic	0.1	(0.1)	0.7	(0.2)	5.0	(0.5)	15.0	(0.8)	25.0	(0.9)	26.9	(0.9)	19.1	(0.8)	7.2	(0.5)	1.1	(0.2)
Denmark	0.0	(0.0)	0.5	(0.1)	3.5	(0.3)	11.9	(0.5)	23.9	(0.8)	30.1	(0.9)	21.6	(0.8)	7.3	(0.5)	1.1	(0.2)
Estonia	0.0	c	0.3	(0.1)	2.1	(0.2)	8.7	(0.5)	21.2	(0.9)	29.9	(0.9)	24.0	(0.8)	11.1	(0.6)	2.8	(0.3)
Finland	0.0	(0.0)	0.8	(0.2)	3.3	(0.4)	9.4	(0.6)	19.2	(0.7)	27.6	(0.8)	25.4	(0.8)	11.9	(0.7)	2.4	(0.3)
France	0.0	(0.0)	1.1	(0.2)	5.7	(0.4)	14.0	(0.7)	22.8	(0.8)	26.6	(0.8)	20.5	(0.7)	8.1	(0.6)	1.1	(0.2)
Germany	0.1	(0.1)	1.3	(0.3)	5.7	(0.5)	13.6	(0.8)	21.1	(0.8)	25.4	(0.8)	21.5	(0.9)	9.5	(0.6)	1.8	(0.2)
Greece	0.1	(0.1)	2.1	(0.3)	9.3	(0.7)	19.0	(0.9)	27.3	(0.8)	25.2	(1.0)	13.3	(0.8)	3.3	(0.4)	0.3	(0.1)
Hungary	0.0	(0.1)	1.2	(0.2)	7.0	(0.6)	17.0	(0.8)	25.2	(0.9)	26.3	(0.9)	17.5	(0.8)	5.2	(0.5)	0.5	(0.1)
Iceland	0.1	(0.1)	2.3	(0.3)	8.0	(0.7)	15.9	(0.8)	24.6	(0.9)	25.1	(0.8)	16.9	(0.7)	6.2	(0.6)	0.9	(0.2)
Ireland	0.0	(0.0)	0.2	(0.1)	2.1	(0.3)	9.5	(0.6)	21.7	(0.8)	30.3	(0.9)	24.1	(0.8)	10.3	(0.6)	1.8	(0.3)
Israel	0.7	(0.2)	5.0	(0.5)	10.4	(0.7)	15.0	(0.9)	19.4	(0.7)	21.6	(0.8)	17.5	(0.8)	8.4	(0.6)	2.0	(0.3)
Italy	0.1	(0.1)	1.7	(0.3)	6.7	(0.6)	14.8	(0.7)	26.3	(0.9)	28.2	(0.9)	16.9	(0.7)	4.9	(0.4)	0.5	(0.1)
Japan	0.1	(0.0)	0.7	(0.2)	4.1	(0.4)	12.0	(0.7)	22.5	(0.9)	28.6	(1.0)	21.9	(0.8)	8.6	(0.6)	1.7	(0.3)
Korea	0.1	(0.1)	1.1	(0.2)	4.3	(0.4)	9.6	(0.7)	19.6	(0.7)	27.6	(0.8)	24.6	(0.8)	10.8	(0.6)	2.3	(0.4)
Latvia	0.0	(0.0)	0.6	(0.1)	5.2	(0.4)	16.6	(0.6)	27.4	(0.8)	28.8	(0.8)	16.6	(0.7)	4.4	(0.4)	0.4	(0.1)
Lithuania	0.0	(0.0)	1.0	(0.2)	6.3	(0.4)	17.0	(0.6)	26.1	(0.8)	27.7	(0.7)	16.9	(0.6)	4.5	(0.4)	0.4	(0.1)
Luxembourg	0.2	(0.1)	2.4	(0.2)	9.2	(0.4)	17.6	(0.6)	23.7	(0.7)	23.5	(0.7)	15.9	(0.6)	6.4	(0.4)	1.3	(0.2)
Mexico	0.0	(0.1)	2.5	(0.4)	13.1	(0.8)	29.1	(1.1)	31.7	(1.0)	17.5	(0.9)	5.3	(0.6)	0.7	(0.2)	0.0	(0.0)
Netherlands*	0.1	(0.1)	1.3	(0.2)	7.0	(0.6)	15.6	(0.7)	23.7	(0.8)	24.3	(1.0)	18.8	(0.8)	7.9	(0.6)	1.2	(0.2)
New Zealand	0.1	(0.1)	1.0	(0.2)	5.2	(0.5)	12.7	(0.6)	20.8	(0.7)	24.6	(0.7)	22.5	(0.7)	10.7	(0.6)	2.4	(0.3)
Norway	0.1	(0.1)	1.7	(0.2)	5.6	(0.4)	11.9	(0.6)	21.5	(0.7)	26.4	(0.9)	21.6	(0.8)	9.6	(0.6)	1.6	(0.2)
Poland	0.0	(0.0)	0.5	(0.1)	3.3	(0.3)	10.8	(0.6)	22.4	(0.8)	27.7	(0.8)	23.0	(0.8)	10.1	(0.7)	2.1	(0.3)
Portugal*	0.0	(0.0)	0.9	(0.2)	5.0	(0.5)	14.3	(0.7)	23.3	(0.7)	28.2	(0.8)	21.0	(0.9)	6.5	(0.6)	0.8	(0.2)
Slovak Republic	0.1	(0.1)	2.3	(0.3)	9.2	(0.7)	19.8	(0.8)	26.9	(0.9)	23.5	(0.9)	13.6	(0.7)	4.1	(0.4)	0.5	(0.2)
Slovenia	0.0	(0.1)	0.6	(0.2)	4.3	(0.4)	12.9	(0.5)	24.5	(0.8)	29.5	(0.9)	20.3	(0.7)	6.8	(0.5)	1.0	(0.2)
Spain	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Sweden	0.2	(0.1)	1.5	(0.2)	5.1	(0.5)	11.6	(0.7)	20.6	(0.8)	25.5	(0.8)	22.3	(0.8)	10.9	(0.7)	2.4	(0.3)
Switzerland	0.1	(0.1)	1.3	(0.3)	7.1	(0.6)	15.1	(0.7)	23.4	(0.9)	26.3	(0.8)	18.5	(0.8)	6.9	(0.6)	1.2	(0.2)
Turkey	0.0	(0.0)	0.7	(0.2)	6.3	(0.6)	19.1	(0.7)	30.2	(0.9)	26.9	(1.0)	13.5	(0.6)	3.1	(0.5)	0.2	(0.1)
United Kingdom	0.0	(0.0)	0.8	(0.2)	4.2	(0.4)	12.3	(0.7)	23.0	(0.7)	27.2	(0.7)	21.0	(0.8)	9.5	(0.6)	2.0	(0.2)
United States*	0.1	(0.1)	1.1	(0.2)	5.4	(0.5)	12.7	(0.8)	21.1	(0.8)	24.7	(0.8)	21.4	(0.8)	10.7	(0.7)	2.8	(0.4)
OECD average-36a	0.1	(0.0)	1.4	(0.0)	6.2	(0.1)	15.0	(0.1)	23.7	(0.1)	26.0	(0.1)	18.9	(0.1)	7.4	(0.1)	1.3	(0.0)
OECD total	0.1	(0.0)	1.3	(0.1)	6.7	(0.2)	15.9	(0.3)	24.0	(0.3)	24.8	(0.3)	18.1	(0.3)	7.6	(0.2)	1.6	(0.1)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.1 [2/2] **Percentage of students at each proficiency level in reading**

		All students																	
		Below Level 1c (less than 189.33 score points)		Level 1c (from 189.33 to less than 262.04 score points)		Level 1b (from 262.04 to less than 334.75 score points)		Level 1a (from 334.75 to less than 407.47 score points)		Level 2 (from 407.47 to less than 480.18 score points)		Level 3 (from 480.18 to less than 552.89 score points)		Level 4 (from 552.89 to less than 625.61 score points)		Level 5 (from 625.61 to less than 698.32 score points)		Level 6 (above 698.32 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners	**Albania**	0.1	(0.1)	2.9	(0.3)	16.4	(0.7)	32.8	(0.9)	29.9	(0.8)	14.0	(0.7)	3.5	(0.4)	0.4	(0.1)	0.0	(0.0)
	Argentina	1.3	(0.2)	6.7	(0.6)	17.4	(0.7)	26.7	(0.9)	25.7	(0.8)	16.2	(0.7)	5.3	(0.5)	0.7	(0.2)	0.0	(0.0)
	Baku (Azerbaijan)	0.1	(0.1)	3.7	(0.4)	19.6	(0.8)	37.0	(1.1)	28.6	(0.9)	9.2	(0.6)	1.6	(0.4)	0.1	(0.1)	0.0	(0.0)
	Belarus	0.0	(0.0)	0.8	(0.2)	5.8	(0.5)	16.8	(0.8)	28.7	(0.8)	28.0	(1.0)	16.0	(0.7)	3.7	(0.4)	0.3	(0.1)
	Bosnia and Herzegovina	0.1	(0.1)	2.8	(0.4)	17.5	(1.0)	33.2	(1.1)	28.8	(1.1)	14.3	(0.9)	3.0	(0.4)	0.2	(0.1)	0.0	c
	Brazil	0.4	(0.1)	5.3	(0.4)	17.7	(0.6)	26.7	(0.7)	24.5	(0.6)	16.3	(0.6)	7.4	(0.5)	1.7	(0.2)	0.2	(0.1)
	Brunei Darussalam	0.3	(0.1)	5.4	(0.3)	19.1	(0.5)	27.0	(0.7)	24.5	(0.6)	15.5	(0.5)	6.9	(0.3)	1.3	(0.2)	0.0	(0.0)
	B-S-J-Z (China)	0.0	(0.0)	0.1	(0.1)	0.7	(0.2)	4.3	(0.5)	14.3	(0.8)	27.9	(1.0)	30.8	(1.0)	17.5	(0.9)	4.2	(0.6)
	Bulgaria	0.3	(0.1)	4.6	(0.6)	17.1	(1.1)	25.1	(0.9)	24.9	(1.0)	17.3	(0.9)	8.4	(0.7)	2.2	(0.3)	0.2	(0.1)
	Costa Rica	0.1	(0.0)	1.8	(0.3)	11.3	(0.7)	28.9	(1.1)	32.1	(1.1)	19.4	(1.1)	5.9	(0.8)	0.6	(0.2)	0.0	c
	Croatia	0.0	(0.0)	0.7	(0.2)	5.0	(0.5)	15.9	(0.8)	28.3	(0.9)	29.0	(1.0)	16.4	(0.8)	4.3	(0.4)	0.4	(0.1)
	Cyprus	0.3	(0.1)	4.3	(0.3)	15.0	(0.6)	24.1	(0.8)	26.9	(0.7)	19.3	(0.6)	8.4	(0.4)	1.7	(0.2)	0.1	(0.1)
	Dominican Republic	1.1	(0.3)	15.9	(0.9)	33.3	(1.1)	28.8	(1.0)	15.0	(0.9)	4.9	(0.5)	0.9	(0.2)	0.1	(0.1)	0.0	(0.0)
	Georgia	0.4	(0.1)	7.0	(0.5)	24.2	(0.9)	32.8	(0.8)	22.9	(0.8)	10.1	(0.6)	2.4	(0.3)	0.2	(0.1)	0.0	(0.0)
	Hong Kong (China)*	0.1	(0.1)	0.9	(0.2)	3.5	(0.4)	8.1	(0.6)	17.8	(0.7)	27.7	(0.7)	27.1	(0.8)	12.5	(0.6)	2.3	(0.3)
	Indonesia	0.2	(0.1)	6.3	(0.6)	26.7	(1.0)	36.7	(1.1)	21.8	(1.0)	7.2	(0.8)	1.1	(0.2)	0.1	(0.0)	0.0	(0.0)
	Jordan	1.1	(0.2)	4.0	(0.5)	11.1	(0.7)	25.0	(0.8)	33.8	(1.0)	20.5	(0.9)	4.3	(0.5)	0.3	(0.1)	0.0	(0.0)
	Kazakhstan	0.1	(0.0)	3.5	(0.3)	22.2	(0.7)	38.4	(0.7)	23.9	(0.5)	8.9	(0.3)	2.6	(0.2)	0.4	(0.1)	0.0	(0.0)
	Kosovo	0.3	(0.1)	8.7	(0.6)	31.7	(0.8)	38.0	(1.0)	17.5	(0.7)	3.6	(0.3)	0.2	(0.1)	0.0	(0.0)	0.0	c
	Lebanon	6.3	(0.6)	16.9	(1.0)	23.0	(0.9)	21.6	(0.8)	17.4	(0.9)	10.5	(0.7)	3.7	(0.5)	0.7	(0.2)	0.0	(0.0)
	Macao (China)	0.0	(0.0)	0.3	(0.1)	2.2	(0.2)	8.2	(0.6)	19.4	(0.8)	29.8	(0.8)	26.1	(0.7)	11.7	(0.6)	2.1	(0.3)
	Malaysia	0.2	(0.1)	3.6	(0.4)	14.2	(0.8)	27.9	(0.9)	31.4	(1.0)	17.9	(0.9)	4.3	(0.6)	0.5	(0.2)	0.0	(0.0)
	Malta	0.7	(0.2)	4.8	(0.4)	11.9	(0.7)	18.5	(0.9)	23.7	(0.9)	21.7	(0.9)	13.4	(0.9)	4.5	(0.5)	0.9	(0.2)
	Moldova	0.4	(0.1)	3.9	(0.5)	13.5	(0.7)	25.2	(0.8)	28.0	(0.9)	20.8	(0.9)	7.2	(0.6)	1.0	(0.3)	0.0	(0.0)
	Montenegro	0.1	(0.1)	2.8	(0.3)	13.5	(0.5)	28.0	(0.7)	30.5	(0.6)	18.3	(0.6)	6.0	(0.4)	0.8	(0.2)	0.0	(0.0)
	Morocco	0.3	(0.1)	8.8	(0.7)	30.8	(1.3)	33.4	(0.9)	20.6	(1.2)	5.6	(0.5)	0.5	(0.1)	0.0	(0.0)	0.0	c
	North Macedonia	1.6	(0.2)	7.3	(0.5)	18.3	(0.8)	27.9	(1.0)	26.6	(0.8)	14.4	(0.6)	3.5	(0.3)	0.3	(0.2)	0.0	(0.0)
	Panama	1.0	(0.2)	8.4	(0.8)	23.4	(0.9)	31.5	(1.0)	23.0	(0.8)	9.9	(0.9)	2.6	(0.4)	0.2	(0.1)	0.0	(0.0)
	Peru	0.4	(0.1)	5.5	(0.5)	19.6	(0.9)	28.9	(0.9)	25.8	(0.7)	14.3	(0.7)	4.8	(0.5)	0.7	(0.2)	0.0	(0.0)
	Philippines	0.5	(0.1)	15.1	(0.9)	38.3	(1.1)	26.7	(0.8)	13.1	(0.7)	5.1	(0.7)	1.1	(0.3)	0.1	(0.0)	0.0	(0.0)
	Qatar	1.2	(0.1)	8.5	(0.3)	17.6	(0.4)	23.6	(0.5)	23.4	(0.4)	15.8	(0.4)	7.3	(0.3)	2.2	(0.2)	0.4	(0.1)
	Romania	0.8	(0.3)	4.3	(0.6)	12.9	(1.0)	22.8	(1.2)	28.1	(1.1)	20.9	(1.3)	8.7	(1.0)	1.3	(0.3)	0.1	(0.1)
	Russia	0.0	(0.0)	1.0	(0.2)	5.6	(0.6)	15.5	(0.9)	28.1	(0.8)	28.0	(0.8)	16.4	(0.7)	4.8	(0.5)	0.6	(0.1)
	Saudi Arabia	0.5	(0.2)	5.3	(0.6)	17.0	(0.9)	29.4	(0.9)	30.4	(1.1)	14.6	(0.8)	2.6	(0.3)	0.1	(0.1)	0.0	c
	Serbia	0.1	(0.1)	2.7	(0.4)	12.2	(0.8)	22.7	(0.8)	27.8	(0.8)	21.8	(0.8)	10.1	(0.7)	2.4	(0.3)	0.2	(0.1)
	Singapore	0.0	(0.0)	0.5	(0.1)	3.0	(0.3)	7.7	(0.4)	14.2	(0.5)	22.3	(0.7)	26.4	(0.6)	18.5	(0.7)	7.3	(0.4)
	Chinese Taipei	0.1	(0.1)	1.2	(0.2)	4.5	(0.4)	12.0	(0.6)	21.8	(0.7)	27.4	(0.8)	22.0	(0.9)	9.3	(0.7)	1.6	(0.3)
	Thailand	0.1	(0.1)	3.6	(0.5)	20.6	(1.1)	35.3	(1.1)	26.0	(1.0)	11.6	(0.9)	2.7	(0.4)	0.2	(0.1)	0.0	(0.0)
	Ukraine	0.2	(0.1)	1.8	(0.3)	7.2	(0.7)	16.7	(0.9)	27.7	(0.8)	28.5	(1.0)	14.5	(0.8)	3.2	(0.4)	0.2	(0.1)
	United Arab Emirates	0.6	(0.1)	5.8	(0.3)	14.9	(0.5)	21.6	(0.4)	23.4	(0.5)	18.1	(0.5)	10.8	(0.6)	4.1	(0.3)	0.7	(0.1)
	Uruguay	0.3	(0.1)	4.0	(0.4)	13.6	(0.8)	24.0	(0.9)	28.1	(1.1)	20.1	(0.8)	8.3	(0.7)	1.5	(0.2)	0.1	(0.1)
	Viet Nam**	0.0	(0.0)	0.1	(0.1)	1.1	(0.3)	8.3	(0.9)	26.9	(1.3)	38.1	(1.2)	20.5	(1.3)	4.6	(0.7)	0.3	(0.1)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.2 [1/2] **Percentage of students at each proficiency level in mathematics**

		All students													
		Below Level 1 (below 357.77 score points)		Level 1 (from 357.77 to less than 420.07 score points)		Level 2 (from 420.07 to less than 482.38 score points)		Level 3 (from 482.38 to less than 544.68 score points)		Level 4 (from 544.68 to less than 606.99 score points)		Level 5 (from 606.99 to less than 669.30 score points)		Level 6 (above 669.30 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	Australia	7.6	(0.5)	14.8	(0.5)	23.4	(0.5)	25.6	(0.5)	18.2	(0.5)	8.0	(0.4)	2.5	(0.3)
	Austria	7.3	(0.7)	13.8	(0.8)	20.8	(1.0)	24.9	(0.9)	20.6	(0.8)	10.0	(0.7)	2.5	(0.3)
	Belgium	6.9	(0.7)	12.8	(0.6)	18.6	(0.7)	23.8	(0.8)	22.2	(0.7)	12.5	(0.6)	3.2	(0.4)
	Canada	5.0	(0.4)	11.3	(0.5)	20.8	(0.6)	25.9	(0.6)	21.7	(0.7)	11.3	(0.5)	4.0	(0.3)
	Chile	24.7	(1.1)	27.2	(0.9)	25.5	(0.9)	15.6	(0.8)	5.7	(0.5)	1.1	(0.2)	0.1	(0.0)
	Colombia	35.5	(1.7)	29.9	(1.2)	21.1	(0.9)	10.0	(0.7)	3.1	(0.4)	0.5	(0.1)	0.0	(0.0)
	Czech Republic	6.6	(0.7)	13.8	(0.7)	22.1	(0.8)	25.2	(0.9)	19.6	(0.7)	9.5	(0.5)	3.1	(0.3)
	Denmark	3.7	(0.4)	10.9	(0.6)	22.0	(0.9)	28.8	(0.8)	23.0	(0.8)	9.5	(0.6)	2.1	(0.3)
	Estonia	2.1	(0.3)	8.1	(0.6)	20.8	(0.8)	29.0	(0.8)	24.6	(0.8)	11.8	(0.7)	3.7	(0.4)
	Finland	3.8	(0.4)	11.1	(0.6)	22.3	(0.9)	28.9	(1.0)	22.7	(0.8)	9.3	(0.5)	1.8	(0.3)
	France	8.0	(0.5)	13.2	(0.6)	21.1	(0.8)	25.6	(0.8)	21.0	(0.8)	9.2	(0.6)	1.8	(0.3)
	Germany	7.6	(0.7)	13.5	(0.8)	20.7	(0.9)	24.0	(0.8)	20.8	(0.8)	10.5	(0.7)	2.8	(0.3)
	Greece	15.3	(1.1)	20.5	(0.9)	26.8	(0.9)	22.5	(1.0)	11.1	(0.6)	3.2	(0.4)	0.5	(0.2)
	Hungary	9.6	(0.7)	16.1	(0.8)	23.6	(0.9)	25.2	(1.0)	17.5	(0.8)	6.5	(0.5)	1.4	(0.3)
	Iceland	7.4	(0.5)	13.3	(0.7)	22.0	(1.0)	26.7	(1.0)	20.2	(0.9)	8.5	(0.6)	1.9	(0.3)
	Ireland	3.8	(0.5)	11.9	(0.7)	24.7	(0.8)	30.5	(0.8)	20.8	(0.8)	7.2	(0.6)	1.0	(0.2)
	Israel	17.7	(1.1)	16.4	(0.8)	20.7	(0.7)	21.0	(0.8)	15.4	(0.8)	7.0	(0.6)	1.8	(0.3)
	Italy	9.1	(0.8)	14.8	(0.9)	22.9	(1.0)	25.6	(0.9)	18.1	(0.8)	7.5	(0.6)	2.0	(0.3)
	Japan	2.9	(0.4)	8.6	(0.6)	18.7	(0.8)	26.4	(0.9)	25.1	(1.0)	14.0	(0.8)	4.3	(0.5)
	Korea	5.4	(0.5)	9.6	(0.6)	17.3	(0.8)	23.4	(0.7)	22.9	(0.8)	14.4	(0.7)	6.9	(0.8)
	Latvia	4.4	(0.5)	12.9	(0.8)	25.8	(0.9)	29.4	(1.0)	19.0	(0.8)	7.1	(0.5)	1.4	(0.2)
	Lithuania	9.3	(0.6)	16.4	(0.7)	24.2	(0.7)	25.2	(0.9)	16.5	(0.8)	6.8	(0.5)	1.7	(0.2)
	Luxembourg	10.9	(0.6)	16.4	(0.6)	21.7	(0.8)	22.6	(0.7)	17.7	(0.7)	8.6	(0.5)	2.3	(0.3)
	Mexico	26.0	(1.2)	30.3	(0.9)	26.4	(0.9)	13.1	(0.8)	3.7	(0.5)	0.5	(0.1)	0.0	(0.0)
	Netherlands*	4.5	(0.6)	11.2	(0.7)	19.0	(1.0)	23.2	(1.1)	23.6	(0.9)	14.2	(0.8)	4.3	(0.5)
	New Zealand	7.6	(0.5)	14.2	(0.6)	22.8	(0.8)	25.0	(0.7)	18.9	(0.7)	8.8	(0.4)	2.7	(0.3)
	Norway	6.5	(0.5)	12.4	(0.6)	21.8	(0.8)	26.5	(0.8)	20.6	(0.9)	9.8	(0.6)	2.4	(0.4)
	Poland	4.2	(0.5)	10.5	(0.6)	20.7	(0.8)	26.5	(0.8)	22.3	(0.7)	11.7	(0.7)	4.1	(0.5)
	Portugal*	9.3	(0.6)	14.0	(0.8)	20.9	(0.8)	24.5	(1.1)	19.7	(0.8)	9.1	(0.6)	2.5	(0.3)
	Slovak Republic	10.7	(0.9)	14.4	(0.6)	21.4	(0.9)	24.2	(0.9)	18.6	(0.9)	8.4	(0.6)	2.3	(0.3)
	Slovenia	4.8	(0.6)	11.7	(0.7)	21.6	(0.9)	26.4	(0.9)	22.0	(0.8)	10.5	(0.8)	3.1	(0.4)
	Spain	8.7	(0.4)	16.0	(0.5)	24.4	(0.4)	26.0	(0.6)	17.5	(0.5)	6.2	(0.3)	1.1	(0.1)
	Sweden	6.0	(0.6)	12.8	(0.8)	21.9	(0.9)	25.7	(0.8)	21.0	(0.8)	10.0	(0.7)	2.6	(0.3)
	Switzerland	4.8	(0.4)	12.0	(0.8)	19.5	(0.9)	24.4	(1.0)	22.3	(0.9)	12.1	(0.7)	4.9	(0.5)
	Turkey	13.8	(0.9)	22.9	(0.8)	27.3	(0.8)	20.4	(0.8)	10.9	(0.5)	3.9	(0.4)	0.9	(0.3)
	United Kingdom	6.4	(0.5)	12.8	(0.6)	22.0	(0.8)	25.5	(0.7)	20.4	(0.7)	9.8	(0.6)	3.1	(0.4)
	United States*	10.2	(0.8)	16.9	(0.9)	24.2	(1.0)	24.1	(1.0)	16.3	(0.9)	6.8	(0.7)	1.5	(0.3)
	OECD average	9.1	(0.1)	14.8	(0.1)	22.2	(0.1)	24.4	(0.1)	18.5	(0.1)	8.5	(0.1)	2.4	(0.1)
	OECD total	11.5	(0.3)	17.2	(0.3)	22.9	(0.3)	22.6	(0.3)	16.3	(0.3)	7.4	(0.2)	2.0	(0.1)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.2 [2/2] **Percentage of students at each proficiency level in mathematics**

	All students													
	Below Level 1 (below 357.77 score points)		Level 1 (from 357.77 to less than 420.07 score points)		Level 2 (from 420.07 to less than 482.38 score points)		Level 3 (from 482.38 to less than 544.68 score points)		Level 4 (from 544.68 to less than 606.99 score points)		Level 5 (from 606.99 to less than 669.30 score points)		Level 6 (above 669.30 score points)	
Partners	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Albania	16.9	(0.9)	25.5	(0.9)	28.6	(1.0)	19.3	(0.8)	7.5	(0.7)	2.0	(0.2)	0.3	(0.1)
Argentina	40.5	(1.6)	28.5	(1.0)	19.6	(0.9)	8.8	(0.7)	2.3	(0.3)	0.3	(0.1)	0.0	(0.0)
Baku (Azerbaijan)	24.7	(1.0)	26.1	(0.8)	25.2	(0.9)	15.7	(0.7)	6.4	(0.6)	1.7	(0.3)	0.3	(0.1)
Belarus	11.4	(0.7)	18.0	(0.7)	24.7	(0.9)	23.4	(0.7)	15.2	(0.7)	6.1	(0.5)	1.2	(0.2)
Bosnia and Herzegovina	28.7	(1.3)	28.9	(1.0)	24.2	(0.9)	13.1	(0.8)	4.3	(0.5)	0.7	(0.2)	0.1	(0.0)
Brazil	41.0	(1.0)	27.1	(0.7)	18.2	(0.7)	9.3	(0.5)	3.4	(0.3)	0.8	(0.2)	0.1	(0.0)
Brunei Darussalam	22.1	(0.8)	25.7	(0.8)	24.0	(0.6)	16.2	(0.5)	8.9	(0.5)	2.7	(0.3)	0.4	(0.1)
B-S-J-Z (China)	0.5	(0.1)	1.9	(0.3)	6.9	(0.5)	17.5	(0.8)	28.9	(1.0)	27.8	(1.0)	16.5	(1.1)
Bulgaria	21.9	(1.4)	22.5	(0.8)	23.7	(1.0)	18.2	(1.0)	9.4	(0.7)	3.3	(0.5)	0.9	(0.2)
Costa Rica	27.8	(1.3)	32.2	(1.2)	25.6	(1.2)	11.2	(1.0)	2.8	(0.5)	0.3	(0.1)	0.0	(0.0)
Croatia	11.0	(0.8)	20.2	(0.8)	27.4	(0.9)	23.3	(0.8)	13.0	(0.8)	4.3	(0.5)	0.8	(0.2)
Cyprus	17.2	(0.6)	19.7	(0.7)	24.7	(0.9)	22.0	(0.8)	12.1	(0.5)	3.7	(0.4)	0.7	(0.1)
Dominican Republic	69.3	(1.4)	21.3	(1.0)	7.3	(0.6)	1.8	(0.4)	0.3	(0.1)	0.0	(0.0)	0.0	c
Georgia	33.7	(1.2)	27.3	(1.1)	21.6	(0.8)	11.9	(0.8)	4.4	(0.5)	0.9	(0.3)	0.1	(0.1)
Hong Kong (China)*	2.8	(0.4)	6.4	(0.6)	13.5	(0.7)	22.1	(0.7)	26.3	(0.9)	19.5	(0.8)	9.5	(0.8)
Indonesia	40.6	(1.6)	31.3	(1.2)	18.6	(1.0)	6.8	(0.7)	2.3	(0.5)	0.4	(0.2)	0.0	(0.0)
Jordan	30.7	(1.4)	28.6	(0.8)	24.0	(0.9)	12.4	(0.8)	3.6	(0.5)	0.6	(0.2)	0.1	(0.1)
Kazakhstan	22.3	(0.8)	26.8	(0.6)	26.6	(0.6)	16.0	(0.6)	6.3	(0.4)	1.6	(0.2)	0.3	(0.1)
Kosovo	47.0	(1.0)	29.6	(1.1)	16.5	(0.8)	5.4	(0.4)	1.4	(0.2)	0.1	(0.1)	0.0	(0.0)
Lebanon	38.0	(1.7)	21.8	(1.0)	19.1	(1.1)	13.1	(0.9)	6.0	(0.5)	1.7	(0.3)	0.3	(0.1)
Macao (China)	1.0	(0.2)	4.0	(0.4)	12.3	(0.8)	24.8	(0.9)	30.3	(1.2)	20.0	(0.8)	7.7	(0.6)
Malaysia	16.1	(0.9)	25.4	(1.0)	28.3	(0.9)	19.3	(0.9)	8.5	(0.7)	2.2	(0.4)	0.3	(0.1)
Malta	14.3	(0.7)	15.9	(0.8)	21.5	(1.0)	23.2	(1.1)	16.6	(0.7)	6.7	(0.6)	1.8	(0.3)
Moldova	26.1	(0.9)	24.2	(0.9)	23.5	(0.9)	16.5	(0.7)	7.3	(0.6)	2.0	(0.3)	0.4	(0.1)
Montenegro	19.9	(0.7)	26.3	(0.7)	27.3	(0.7)	17.9	(0.5)	6.9	(0.4)	1.6	(0.2)	0.2	(0.1)
Morocco	47.1	(1.9)	28.5	(1.0)	16.9	(1.0)	6.2	(0.6)	1.2	(0.2)	0.1	(0.1)	0.0	(0.0)
North Macedonia	35.2	(0.8)	25.8	(0.8)	21.3	(0.7)	12.1	(0.7)	4.5	(0.4)	1.0	(0.2)	0.1	(0.1)
Panama	53.7	(1.4)	27.5	(1.0)	13.5	(0.8)	4.3	(0.6)	0.9	(0.2)	0.1	(0.1)	0.0	(0.0)
Peru	32.0	(1.2)	28.3	(0.8)	23.1	(0.9)	11.6	(0.7)	4.1	(0.5)	0.8	(0.2)	0.1	(0.0)
Philippines	54.4	(1.7)	26.3	(0.9)	13.6	(1.0)	4.7	(0.7)	0.9	(0.3)	0.1	(0.1)	0.0	(0.0)
Qatar	29.7	(0.7)	24.0	(0.5)	21.9	(0.5)	14.6	(0.4)	6.9	(0.3)	2.4	(0.2)	0.6	(0.1)
Romania	22.6	(1.6)	23.9	(1.2)	24.5	(1.1)	17.3	(1.1)	8.5	(1.0)	2.7	(0.5)	0.4	(0.2)
Russia	6.8	(0.7)	14.9	(0.8)	25.0	(0.9)	27.5	(0.9)	17.8	(0.8)	6.6	(0.6)	1.5	(0.2)
Saudi Arabia	42.8	(1.6)	29.9	(1.0)	18.8	(1.1)	6.8	(0.6)	1.5	(0.3)	0.2	(0.1)	0.0	(0.0)
Serbia	18.1	(1.1)	21.6	(0.8)	24.1	(0.8)	19.2	(0.8)	11.7	(0.7)	4.2	(0.4)	1.0	(0.2)
Singapore	1.8	(0.2)	5.3	(0.4)	11.1	(0.5)	19.1	(0.7)	25.8	(0.8)	23.2	(0.7)	13.8	(0.8)
Chinese Taipei	5.0	(0.4)	9.0	(0.5)	16.1	(0.7)	23.2	(0.8)	23.5	(0.8)	15.6	(0.8)	7.6	(0.8)
Thailand	25.0	(1.3)	27.7	(1.0)	24.6	(1.0)	14.3	(0.8)	6.1	(0.7)	1.9	(0.3)	0.3	(0.1)
Ukraine	15.6	(1.2)	20.3	(1.0)	26.2	(1.0)	21.5	(1.0)	11.5	(0.8)	4.0	(0.5)	1.0	(0.3)
United Arab Emirates	24.2	(0.9)	21.3	(0.6)	21.5	(0.5)	17.2	(0.6)	10.4	(0.5)	4.2	(0.3)	1.2	(0.1)
Uruguay	24.6	(1.1)	26.1	(1.3)	26.5	(1.0)	15.8	(1.0)	6.0	(0.6)	1.0	(0.2)	0.1	(0.0)
Viet Nam**	3.1	(0.6)	12.6	(1.2)	27.3	(1.3)	31.4	(1.1)	18.8	(1.2)	5.7	(0.7)	1.1	(0.3)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.3 [1/2] **Percentage of students at each proficiency level in science**

	All students															
	Below Level 1b (below 260.54 score points)		Level 1b (from 260.54 to less than 334.94 score points)		Level 1a (from 334.94 to less than 409.54 score points)		Level 2 (from 409.54 to less than 484.14 score points)		Level 3 (from 484.14 to less than 558.73 score points)		Level 4 (from 558.73 to less than 633.33 score points)		Level 5 (from 633.33 to less than 707.93 score points)		Level 6 (above 707.93 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD																
Australia	0.6	(0.1)	4.5	(0.3)	13.7	(0.5)	23.0	(0.6)	27.5	(0.6)	21.2	(0.6)	7.9	(0.4)	1.6	(0.2)
Austria	0.6	(0.2)	4.8	(0.5)	16.5	(0.9)	25.0	(0.8)	27.6	(0.8)	19.2	(0.8)	5.8	(0.6)	0.5	(0.1)
Belgium	0.6	(0.1)	5.3	(0.5)	14.2	(0.6)	22.2	(0.7)	28.4	(0.8)	21.3	(0.7)	7.3	(0.4)	0.7	(0.2)
Canada	0.4	(0.1)	2.6	(0.2)	10.5	(0.4)	22.4	(0.6)	29.3	(0.6)	23.5	(0.7)	9.5	(0.5)	1.8	(0.2)
Chile	1.0	(0.2)	8.8	(0.7)	25.5	(1.0)	33.1	(1.0)	22.6	(1.0)	7.9	(0.6)	1.0	(0.2)	0.0	(0.0)
Colombia	2.1	(0.3)	15.3	(1.1)	33.0	(1.1)	29.6	(1.2)	15.4	(0.8)	4.2	(0.4)	0.4	(0.1)	0.0	(0.0)
Czech Republic	0.4	(0.1)	3.9	(0.4)	14.5	(0.8)	25.9	(1.0)	28.7	(1.0)	19.1	(0.8)	6.6	(0.5)	1.0	(0.2)
Denmark	0.7	(0.2)	4.1	(0.3)	13.9	(0.6)	26.6	(0.7)	30.1	(0.9)	19.1	(0.8)	5.0	(0.5)	0.5	(0.2)
Estonia	0.1	(0.1)	1.1	(0.2)	7.5	(0.5)	21.5	(0.7)	32.1	(0.9)	25.4	(0.8)	10.2	(0.5)	2.0	(0.2)
Finland	0.4	(0.1)	2.8	(0.3)	9.7	(0.6)	21.1	(0.7)	28.9	(0.8)	24.9	(0.8)	10.5	(0.6)	1.8	(0.3)
France	0.6	(0.2)	5.0	(0.4)	14.9	(0.8)	24.6	(0.9)	28.3	(0.7)	20.0	(0.9)	5.9	(0.5)	0.6	(0.1)
Germany	0.8	(0.2)	5.0	(0.5)	13.8	(0.7)	22.0	(0.9)	26.9	(0.9)	21.5	(1.0)	8.5	(0.6)	1.5	(0.2)
Greece	1.2	(0.3)	8.1	(0.8)	22.4	(1.0)	31.6	(0.9)	26.0	(1.0)	9.3	(0.6)	1.3	(0.2)	0.0	(0.0)
Hungary	0.6	(0.2)	5.7	(0.6)	17.8	(0.9)	26.1	(1.0)	28.1	(0.9)	17.0	(0.7)	4.3	(0.5)	0.4	(0.1)
Iceland	0.5	(0.2)	5.9	(0.5)	18.6	(0.8)	28.3	(0.9)	27.7	(1.0)	15.2	(0.8)	3.6	(0.4)	0.2	(0.1)
Ireland	0.3	(0.1)	3.3	(0.3)	13.4	(0.7)	26.9	(0.9)	31.3	(0.9)	19.0	(0.7)	5.4	(0.5)	0.5	(0.2)
Israel	3.2	(0.4)	10.7	(0.7)	19.2	(0.9)	23.1	(0.9)	22.9	(0.8)	15.1	(0.8)	5.2	(0.4)	0.7	(0.1)
Italy	1.1	(0.2)	6.6	(0.5)	18.2	(0.9)	30.2	(1.0)	27.8	(1.1)	13.4	(0.7)	2.6	(0.4)	0.2	(0.1)
Japan	0.2	(0.1)	1.8	(0.3)	8.9	(0.6)	19.9	(0.8)	29.7	(1.1)	26.5	(0.9)	11.4	(0.7)	1.6	(0.3)
Korea	0.5	(0.1)	3.1	(0.3)	10.6	(0.7)	21.0	(0.8)	28.6	(0.9)	24.5	(0.9)	10.0	(0.6)	1.8	(0.3)
Latvia	0.3	(0.1)	3.4	(0.4)	14.8	(0.7)	29.5	(0.8)	31.5	(1.1)	16.8	(0.8)	3.5	(0.4)	0.3	(0.1)
Lithuania	0.5	(0.2)	4.7	(0.4)	17.0	(0.8)	28.4	(0.8)	28.7	(0.8)	16.3	(0.6)	4.0	(0.3)	0.5	(0.1)
Luxembourg	0.8	(0.2)	6.8	(0.4)	19.2	(0.6)	25.7	(0.8)	25.6	(0.8)	16.6	(0.6)	4.9	(0.5)	0.5	(0.2)
Mexico	1.0	(0.3)	11.6	(1.0)	34.2	(1.3)	33.9	(0.9)	15.5	(0.9)	3.5	(0.5)	0.3	(0.1)	0.0	c
Netherlands*	0.9	(0.2)	4.8	(0.5)	14.4	(0.8)	22.4	(0.8)	24.9	(1.1)	22.1	(1.0)	9.1	(0.7)	1.5	(0.3)
New Zealand	0.6	(0.2)	4.3	(0.4)	13.1	(0.6)	22.0	(0.6)	26.8	(0.7)	21.8	(0.7)	9.5	(0.6)	1.8	(0.3)
Norway	1.1	(0.2)	5.7	(0.4)	14.1	(0.8)	25.0	(0.9)	28.6	(0.7)	18.7	(0.7)	6.1	(0.5)	0.7	(0.1)
Poland	0.2	(0.1)	2.5	(0.3)	11.1	(0.7)	24.9	(0.8)	30.0	(1.0)	22.0	(0.8)	8.1	(0.7)	1.2	(0.2)
Portugal*	0.4	(0.1)	4.4	(0.6)	14.7	(0.9)	26.2	(0.9)	29.4	(1.0)	19.2	(0.9)	5.1	(0.5)	0.5	(0.2)
Slovak Republic	1.4	(0.2)	7.9	(0.6)	19.9	(0.7)	28.5	(0.9)	25.3	(0.8)	13.2	(0.6)	3.4	(0.3)	0.3	(0.1)
Slovenia	0.2	(0.1)	2.5	(0.3)	11.9	(0.6)	24.6	(0.8)	31.8	(1.0)	21.8	(0.9)	6.7	(0.5)	0.6	(0.2)
Spain	0.6	(0.1)	4.5	(0.3)	16.2	(0.5)	28.4	(0.5)	29.4	(0.5)	16.8	(0.4)	3.9	(0.2)	0.3	(0.1)
Sweden	0.6	(0.2)	4.6	(0.5)	13.8	(0.7)	24.0	(0.7)	28.0	(0.8)	20.7	(0.9)	7.3	(0.5)	1.0	(0.2)
Switzerland	0.4	(0.1)	4.6	(0.5)	15.2	(0.8)	24.9	(0.9)	27.8	(0.9)	19.3	(1.0)	6.9	(0.7)	0.9	(0.2)
Turkey	0.3	(0.1)	4.7	(0.4)	20.1	(0.8)	32.8	(1.0)	27.3	(1.0)	12.3	(0.7)	2.3	(0.4)	0.1	(0.1)
United Kingdom	0.6	(0.2)	3.9	(0.4)	12.9	(0.6)	24.0	(0.8)	28.1	(0.8)	20.8	(0.7)	8.2	(0.6)	1.5	(0.2)
United States*	0.5	(0.2)	4.4	(0.5)	13.7	(0.8)	23.6	(0.9)	27.5	(0.9)	21.1	(0.9)	7.9	(0.7)	1.3	(0.2)
OECD average	0.7	(0.0)	5.2	(0.1)	16.0	(0.1)	25.8	(0.1)	27.4	(0.1)	18.1	(0.1)	5.9	(0.1)	0.8	(0.0)
OECD total	0.7	(0.1)	5.6	(0.2)	17.2	(0.3)	25.9	(0.3)	26.0	(0.3)	17.6	(0.3)	6.1	(0.2)	0.9	(0.1)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.3 [2/2] **Percentage of students at each proficiency level in science**

	All students															
	Below Level 1b (below 260.54 score points)		Level 1b (from 260.54 to less than 334.94 score points)		Level 1a (from 334.94 to less than 409.54 score points)		Level 2 (from 409.54 to less than 484.14 score points)		Level 3 (from 484.14 to less than 558.73 score points)		Level 4 (from 558.73 to less than 633.33 score points)		Level 5 (from 633.33 to less than 707.93 score points)		Level 6 (above 707.93 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners																
Albania	1.5	(0.2)	11.7	(0.7)	33.7	(1.0)	34.8	(1.1)	15.1	(0.7)	2.9	(0.3)	0.2	(0.1)	0.0	(0.0)
Argentina	4.9	(0.6)	18.2	(1.0)	30.4	(1.1)	27.0	(0.9)	15.0	(0.8)	4.1	(0.4)	0.5	(0.1)	0.0	(0.0)
Baku (Azerbaijan)	2.5	(0.3)	17.3	(1.0)	38.0	(1.0)	29.9	(0.9)	10.3	(0.7)	1.8	(0.4)	0.1	(0.1)	0.0	c
Belarus	0.5	(0.2)	5.0	(0.5)	18.7	(0.9)	31.3	(0.9)	28.8	(0.8)	13.1	(0.8)	2.5	(0.4)	0.1	(0.1)
Bosnia and Herzegovina	2.9	(0.4)	18.2	(0.9)	35.6	(1.0)	29.4	(1.2)	11.7	(0.9)	1.9	(0.3)	0.1	(0.1)	0.0	c
Brazil	4.0	(0.4)	19.9	(0.7)	31.4	(0.8)	25.3	(0.7)	13.9	(0.7)	4.6	(0.4)	0.8	(0.1)	0.0	(0.0)
Brunei Darussalam	1.9	(0.3)	14.2	(0.6)	29.7	(0.8)	25.5	(0.5)	17.4	(0.5)	9.0	(0.4)	2.1	(0.3)	0.1	(0.1)
B-S-J-Z (China)	0.0	(0.0)	0.3	(0.1)	1.8	(0.3)	8.4	(0.6)	23.4	(0.9)	34.6	(1.0)	24.3	(1.1)	7.2	(0.7)
Bulgaria	3.0	(0.5)	15.3	(1.0)	28.3	(0.9)	26.7	(1.1)	17.9	(0.9)	7.4	(0.6)	1.4	(0.3)	0.1	(0.1)
Costa Rica	1.3	(0.3)	12.0	(0.8)	34.5	(1.2)	34.4	(1.2)	14.9	(1.2)	2.8	(0.6)	0.1	(0.1)	0.0	c
Croatia	0.6	(0.2)	5.6	(0.5)	19.1	(0.9)	30.0	(0.8)	26.9	(0.9)	14.2	(0.7)	3.3	(0.4)	0.3	(0.1)
Cyprus	2.0	(0.3)	11.9	(0.6)	25.0	(0.8)	28.9	(1.0)	21.4	(0.7)	9.1	(0.4)	1.5	(0.2)	0.1	(0.1)
Dominican Republic	13.6	(1.0)	39.6	(1.3)	31.6	(1.3)	12.3	(0.9)	2.6	(0.4)	0.3	(0.1)	0.0	(0.0)	0.0	c
Georgia	5.8	(0.5)	22.9	(0.9)	35.7	(0.9)	24.3	(0.9)	9.5	(0.6)	1.7	(0.3)	0.1	(0.1)	0.0	c
Hong Kong (China)*	0.2	(0.1)	2.4	(0.3)	8.9	(0.6)	21.7	(0.8)	33.8	(0.9)	25.0	(0.9)	7.1	(0.6)	0.7	(0.2)
Indonesia	1.8	(0.3)	16.8	(1.0)	41.4	(1.1)	29.2	(1.2)	9.2	(0.8)	1.6	(0.3)	0.1	(0.0)	0.0	(0.0)
Jordan	3.2	(0.4)	11.0	(0.8)	26.2	(0.9)	32.4	(1.0)	20.7	(0.9)	6.0	(0.5)	0.6	(0.2)	0.0	(0.0)
Kazakhstan	2.2	(0.3)	17.8	(0.7)	40.3	(0.8)	26.9	(0.8)	9.9	(0.5)	2.5	(0.3)	0.4	(0.1)	0.0	(0.0)
Kosovo	4.2	(0.4)	29.3	(0.9)	43.1	(1.0)	19.2	(0.7)	3.9	(0.4)	0.4	(0.1)	0.0	(0.0)	0.0	c
Lebanon	8.9	(0.8)	23.6	(1.2)	29.7	(1.0)	21.8	(1.0)	11.8	(0.8)	3.6	(0.4)	0.5	(0.2)	0.0	(0.0)
Macao (China)	0.1	(0.1)	0.8	(0.2)	5.1	(0.5)	17.2	(0.7)	32.3	(1.0)	30.8	(0.9)	11.9	(0.6)	1.7	(0.3)
Malaysia	0.7	(0.2)	8.3	(0.7)	27.6	(1.0)	35.9	(1.0)	21.5	(0.9)	5.4	(0.8)	0.6	(0.2)	0.0	(0.0)
Malta	3.4	(0.4)	10.8	(0.7)	19.4	(0.7)	24.9	(0.9)	23.7	(0.9)	13.5	(0.7)	3.9	(0.4)	0.5	(0.1)
Moldova	2.4	(0.3)	12.7	(0.7)	27.4	(0.9)	29.7	(0.9)	20.2	(0.8)	6.6	(0.5)	0.8	(0.2)	0.0	(0.0)
Montenegro	2.2	(0.3)	14.6	(0.6)	31.4	(0.8)	31.5	(0.7)	15.9	(0.6)	4.0	(0.3)	0.3	(0.1)	0.0	(0.0)
Morocco	2.7	(0.4)	26.1	(1.4)	40.7	(1.1)	24.0	(1.4)	6.1	(0.6)	0.4	(0.1)	0.0	(0.0)	0.0	c
North Macedonia	4.5	(0.4)	15.5	(0.6)	29.4	(0.8)	28.2	(0.9)	16.4	(0.7)	5.2	(0.4)	0.8	(0.2)	0.0	(0.0)
Panama	10.5	(0.9)	27.3	(1.1)	33.5	(1.3)	19.7	(0.8)	7.4	(0.7)	1.5	(0.3)	0.1	(0.1)	0.0	c
Peru	2.7	(0.4)	17.3	(0.9)	34.5	(1.1)	29.0	(0.8)	13.2	(0.8)	3.1	(0.5)	0.2	(0.1)	0.0	(0.0)
Philippines	7.5	(0.8)	35.3	(1.4)	35.2	(1.2)	15.4	(0.8)	5.6	(0.7)	1.0	(0.3)	0.1	(0.0)	0.0	c
Qatar	5.2	(0.3)	16.6	(0.4)	26.5	(0.6)	24.9	(0.5)	17.0	(0.4)	7.5	(0.3)	2.0	(0.2)	0.2	(0.1)
Romania	2.9	(0.5)	13.1	(1.2)	28.0	(1.4)	29.8	(1.0)	18.9	(1.3)	6.4	(0.8)	0.9	(0.2)	0.0	(0.0)
Russia	0.4	(0.2)	4.1	(0.5)	16.7	(0.9)	31.7	(0.9)	30.0	(0.9)	14.0	(0.8)	2.9	(0.4)	0.2	(0.1)
Saudi Arabia	4.9	(0.6)	21.7	(1.0)	35.6	(1.0)	26.6	(1.0)	9.6	(0.7)	1.5	(0.3)	0.1	(0.0)	0.0	c
Serbia	1.9	(0.3)	11.1	(0.8)	25.3	(1.0)	29.9	(0.9)	21.1	(0.9)	9.1	(0.7)	1.5	(0.2)	0.1	(0.0)
Singapore	0.2	(0.1)	1.8	(0.2)	7.1	(0.4)	15.1	(0.7)	25.4	(0.7)	29.7	(0.7)	17.0	(0.5)	3.8	(0.3)
Chinese Taipei	0.7	(0.2)	3.3	(0.3)	11.2	(0.6)	21.1	(0.9)	28.5	(0.9)	23.5	(0.8)	10.0	(0.8)	1.6	(0.3)
Thailand	1.3	(0.3)	11.6	(0.8)	31.6	(1.1)	31.7	(0.9)	17.8	(1.0)	5.3	(0.7)	0.7	(0.2)	0.0	(0.0)
Ukraine	1.0	(0.2)	6.3	(0.6)	19.2	(0.9)	30.0	(1.1)	26.7	(1.1)	13.4	(0.8)	3.2	(0.5)	0.3	(0.1)
United Arab Emirates	3.7	(0.2)	14.4	(0.5)	24.7	(0.6)	25.6	(0.5)	19.2	(0.5)	9.5	(0.5)	2.6	(0.2)	0.3	(0.1)
Uruguay	2.1	(0.4)	13.2	(0.8)	28.6	(1.0)	30.6	(1.0)	18.7	(0.9)	6.1	(0.5)	0.7	(0.2)	0.0	(0.0)
Viet Nam**	0.0	(0.0)	0.3	(0.1)	3.6	(0.5)	18.1	(1.2)	36.2	(1.2)	29.8	(1.1)	10.6	(0.8)	1.5	(0.3)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.4 [1/2] **Mean score and variation in reading performance**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD	Australia	503	(1.6)	109	(0.9)	315	(2.7)	357	(2.8)	429	(2.2)	507	(1.9)	580	(2.0)	640	(2.2)	673	(2.6)
	Austria	484	(2.7)	99	(1.2)	318	(3.9)	350	(3.7)	413	(4.1)	488	(3.8)	558	(2.9)	612	(2.9)	641	(2.9)
	Belgium	493	(2.3)	103	(1.3)	317	(4.0)	352	(3.8)	421	(3.2)	498	(2.7)	568	(2.6)	623	(2.6)	653	(2.8)
	Canada	520	(1.8)	100	(0.8)	349	(2.8)	388	(2.4)	452	(2.3)	524	(2.2)	592	(2.0)	646	(2.3)	677	(2.8)
	Chile	452	(2.6)	92	(1.2)	298	(3.7)	331	(3.6)	389	(3.1)	453	(3.2)	517	(3.4)	572	(3.3)	602	(3.5)
	Colombia	412	(3.3)	89	(1.5)	272	(4.1)	300	(3.7)	350	(3.5)	408	(3.8)	472	(4.1)	532	(4.7)	566	(4.9)
	Czech Republic	490	(2.5)	97	(1.6)	328	(5.2)	362	(4.3)	422	(3.7)	492	(3.0)	560	(2.9)	616	(2.8)	647	(3.1)
	Denmark	501	(1.8)	92	(1.2)	344	(4.0)	380	(3.0)	439	(2.7)	504	(2.2)	566	(2.1)	618	(2.6)	647	(3.3)
	Estonia	523	(1.8)	93	(1.2)	367	(3.8)	402	(3.5)	460	(2.6)	524	(2.3)	587	(2.3)	643	(3.1)	676	(3.7)
	Finland	520	(2.3)	100	(1.3)	345	(4.7)	387	(4.2)	455	(3.2)	527	(2.8)	591	(2.5)	643	(3.0)	672	(3.3)
	France	493	(2.3)	101	(1.5)	319	(4.3)	355	(3.5)	423	(3.0)	497	(3.0)	567	(3.3)	622	(3.6)	651	(4.0)
	Germany	498	(3.0)	106	(1.5)	316	(5.0)	354	(4.5)	424	(4.4)	504	(4.1)	576	(3.5)	632	(3.5)	663	(3.6)
	Greece	457	(3.6)	97	(1.6)	292	(4.8)	326	(4.9)	390	(4.9)	460	(4.1)	526	(3.7)	583	(3.9)	614	(5.0)
	Hungary	476	(2.3)	98	(1.3)	311	(3.7)	346	(4.0)	407	(3.0)	479	(3.1)	547	(2.9)	602	(3.7)	631	(4.1)
	Iceland	474	(1.7)	105	(1.3)	293	(4.4)	332	(4.0)	402	(3.3)	477	(2.7)	549	(3.0)	609	(3.3)	640	(3.8)
	Ireland	518	(2.2)	91	(1.0)	364	(4.1)	398	(3.5)	456	(2.8)	520	(2.4)	583	(2.6)	635	(2.8)	663	(3.8)
	Israel	470	(3.7)	124	(1.9)	256	(5.4)	296	(5.9)	381	(5.8)	479	(4.9)	563	(3.8)	628	(3.7)	663	(3.9)
	Italy	476	(2.4)	97	(1.7)	306	(5.5)	345	(4.6)	413	(3.2)	481	(2.9)	545	(3.0)	598	(3.4)	628	(3.5)
	Japan	504	(2.7)	97	(1.7)	337	(5.1)	374	(4.5)	438	(3.7)	508	(3.0)	572	(3.1)	627	(3.7)	657	(4.1)
	Korea	514	(2.9)	102	(1.7)	329	(5.8)	377	(4.9)	449	(3.8)	522	(3.1)	585	(3.1)	640	(3.9)	669	(4.1)
	Latvia	479	(1.6)	90	(1.1)	328	(3.6)	360	(3.2)	415	(2.3)	480	(2.2)	542	(2.3)	595	(2.7)	624	(3.0)
	Lithuania	476	(1.5)	94	(1.0)	316	(3.5)	351	(2.7)	410	(2.6)	479	(2.3)	543	(1.9)	597	(1.8)	625	(3.2)
	Luxembourg	470	(1.1)	108	(1.0)	291	(3.1)	325	(2.1)	392	(2.0)	472	(1.8)	548	(1.9)	612	(2.8)	646	(3.9)
	Mexico	420	(2.7)	84	(1.6)	286	(3.9)	314	(3.5)	362	(2.8)	419	(2.9)	476	(3.5)	530	(4.2)	562	(5.8)
	Netherlands*	485	(2.7)	105	(1.7)	309	(5.2)	344	(4.4)	410	(3.5)	486	(3.7)	562	(3.4)	621	(3.3)	651	(3.4)
	New Zealand	506	(2.0)	106	(1.3)	322	(4.8)	362	(3.7)	432	(3.2)	511	(2.9)	584	(2.1)	640	(2.9)	671	(2.9)
	Norway	499	(2.2)	106	(1.3)	310	(4.3)	356	(4.3)	430	(3.2)	506	(2.7)	576	(3.1)	632	(2.9)	661	(3.0)
	Poland	512	(2.7)	97	(1.4)	347	(4.5)	384	(3.6)	446	(2.9)	515	(3.3)	581	(3.4)	636	(4.0)	667	(4.1)
	Portugal*	492	(2.4)	96	(1.2)	327	(4.7)	362	(4.0)	425	(3.4)	497	(2.9)	562	(2.9)	613	(2.7)	640	(4.4)
	Slovak Republic	458	(2.2)	100	(1.4)	291	(4.3)	326	(4.0)	388	(3.1)	458	(2.9)	529	(3.1)	590	(3.3)	623	(3.5)
	Slovenia	495	(1.2)	94	(1.2)	335	(3.9)	372	(3.0)	431	(2.2)	499	(1.9)	561	(2.1)	614	(2.8)	644	(3.4)
	Spain	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
	Sweden	506	(3.0)	108	(1.5)	317	(5.5)	360	(5.7)	434	(4.1)	512	(3.4)	583	(3.2)	640	(3.5)	672	(3.7)
	Switzerland	484	(3.1)	103	(1.5)	308	(5.1)	345	(4.6)	413	(4.0)	488	(3.6)	558	(3.8)	615	(4.0)	647	(4.4)
	Turkey	466	(2.2)	88	(1.6)	321	(4.6)	351	(4.1)	404	(3.0)	466	(2.6)	527	(2.4)	581	(3.1)	610	(4.6)
	United Kingdom	504	(2.6)	100	(1.3)	334	(4.4)	372	(4.3)	435	(3.2)	506	(2.7)	575	(3.1)	632	(3.5)	664	(3.8)
	United States*	505	(3.6)	108	(1.6)	321	(5.7)	361	(5.3)	430	(4.4)	510	(4.1)	584	(4.3)	643	(3.9)	676	(4.6)
	OECD average-36a	487	(0.4)	99	(0.2)	318	(0.7)	354	(0.7)	419	(0.6)	490	(0.5)	558	(0.5)	614	(0.5)	644	(0.6)
	OECD total	485	(1.2)	105	(0.6)	311	(1.6)	347	(1.5)	411	(1.4)	486	(1.4)	560	(1.4)	620	(1.6)	654	(1.8)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.4 [2/2] **Mean score and variation in reading performance**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners	**Albania**	405	(1.9)	80	(1.2)	277	(2.9)	303	(2.9)	349	(2.2)	403	(2.1)	459	(2.8)	510	(3.3)	542	(4.1)
	Argentina	402	(3.0)	98	(1.5)	240	(4.5)	274	(4.2)	333	(3.4)	402	(3.6)	471	(3.6)	529	(3.4)	561	(3.9)
	Baku (Azerbaijan)	389	(2.5)	74	(1.7)	270	(2.6)	294	(2.5)	338	(2.4)	389	(2.4)	438	(3.0)	485	(4.6)	514	(6.3)
	Belarus	474	(2.4)	89	(1.3)	322	(4.5)	355	(3.4)	412	(3.1)	475	(3.0)	538	(3.0)	589	(3.1)	617	(4.0)
	Bosnia and Herzegovina	403	(2.9)	79	(1.2)	278	(3.1)	303	(2.8)	346	(3.0)	400	(3.5)	458	(3.7)	509	(4.1)	537	(4.0)
	Brazil	413	(2.1)	100	(1.3)	258	(2.6)	286	(2.6)	340	(2.3)	408	(2.4)	482	(3.1)	548	(3.7)	584	(4.1)
	Brunei Darussalam	408	(0.9)	97	(0.8)	258	(1.9)	284	(1.9)	335	(1.4)	403	(1.5)	476	(1.7)	542	(2.5)	578	(2.5)
	B-S-J-Z (China)	555	(2.7)	87	(1.7)	406	(5.9)	441	(4.2)	498	(3.5)	559	(2.9)	617	(3.1)	666	(3.5)	692	(4.8)
	Bulgaria	420	(3.9)	101	(1.8)	263	(4.3)	290	(4.5)	344	(4.9)	416	(4.8)	491	(5.0)	557	(5.2)	594	(5.3)
	Costa Rica	426	(3.4)	81	(1.7)	295	(3.8)	323	(3.1)	370	(2.9)	424	(3.5)	483	(4.5)	534	(5.9)	563	(6.4)
	Croatia	479	(2.7)	89	(1.7)	329	(5.2)	362	(4.6)	418	(3.7)	480	(3.2)	542	(2.9)	594	(3.2)	623	(3.9)
	Cyprus	424	(1.4)	98	(0.9)	265	(2.7)	295	(2.9)	353	(2.3)	424	(1.9)	494	(2.0)	554	(2.6)	587	(3.0)
	Dominican Republic	342	(2.9)	82	(1.8)	221	(2.8)	241	(2.5)	281	(2.7)	334	(3.2)	395	(4.0)	453	(5.5)	488	(6.1)
	Georgia	380	(2.2)	84	(1.2)	249	(3.1)	274	(2.5)	319	(2.6)	374	(2.7)	436	(2.8)	493	(3.6)	526	(3.8)
	Hong Kong (China)*	524	(2.7)	99	(1.5)	342	(6.7)	390	(5.5)	463	(3.7)	533	(2.9)	595	(2.6)	645	(2.5)	673	(3.3)
	Indonesia	371	(2.6)	75	(1.7)	254	(3.6)	277	(3.1)	318	(2.8)	367	(2.8)	420	(3.6)	472	(5.1)	502	(5.7)
	Jordan	419	(2.9)	87	(1.7)	261	(6.9)	303	(5.7)	366	(3.9)	426	(3.0)	480	(2.6)	524	(3.1)	550	(3.6)
	Kazakhstan	387	(1.5)	77	(1.2)	271	(2.5)	294	(2.2)	333	(1.7)	380	(1.5)	433	(1.9)	490	(2.9)	527	(4.1)
	Kosovo	353	(1.1)	68	(0.7)	245	(2.2)	265	(2.1)	304	(1.9)	352	(1.7)	398	(1.7)	442	(2.0)	470	(3.1)
	Lebanon	353	(4.3)	113	(1.6)	180	(4.9)	211	(4.6)	268	(4.6)	347	(5.7)	434	(5.2)	507	(5.0)	546	(5.7)
	Macao (China)	525	(1.2)	92	(1.1)	365	(5.0)	403	(3.2)	464	(2.3)	530	(1.7)	590	(2.1)	641	(3.0)	670	(2.8)
	Malaysia	415	(2.9)	85	(1.6)	273	(3.5)	302	(3.4)	357	(3.1)	417	(3.2)	474	(3.4)	524	(4.2)	552	(5.0)
	Malta	448	(1.7)	113	(1.2)	258	(4.2)	295	(3.2)	369	(3.0)	452	(2.6)	529	(3.0)	593	(3.3)	628	(4.3)
	Moldova	424	(2.4)	93	(1.6)	268	(4.4)	301	(3.3)	358	(2.9)	425	(3.1)	491	(3.4)	544	(3.7)	573	(4.9)
	Montenegro	421	(1.1)	86	(0.8)	281	(2.6)	310	(2.1)	360	(1.6)	420	(1.7)	480	(1.6)	534	(2.0)	566	(2.7)
	Morocco	359	(3.1)	75	(1.1)	244	(2.6)	265	(2.6)	304	(3.0)	355	(3.9)	412	(4.0)	460	(3.6)	488	(3.9)
	North Macedonia	393	(1.1)	94	(1.0)	233	(3.4)	268	(2.7)	328	(2.2)	395	(1.9)	460	(1.8)	513	(2.4)	543	(2.7)
	Panama	377	(3.0)	88	(1.9)	237	(4.0)	265	(3.7)	315	(3.0)	374	(3.0)	436	(4.2)	493	(5.6)	528	(6.7)
	Peru	401	(3.0)	92	(1.5)	256	(3.5)	283	(2.9)	334	(3.3)	397	(3.3)	463	(3.8)	523	(4.9)	558	(6.3)
	Philippines	340	(3.3)	80	(2.3)	230	(2.6)	248	(2.3)	281	(2.3)	327	(3.1)	388	(4.7)	453	(7.2)	491	(8.3)
	Qatar	407	(0.8)	110	(0.6)	233	(1.9)	264	(1.8)	326	(1.5)	405	(1.3)	483	(1.2)	552	(1.8)	592	(2.1)
	Romania	428	(5.1)	98	(2.2)	261	(6.5)	297	(6.0)	361	(6.1)	431	(6.0)	497	(6.0)	554	(5.9)	584	(5.5)
	Russia	479	(3.1)	93	(1.8)	321	(5.4)	357	(4.8)	416	(3.7)	480	(3.4)	543	(3.3)	597	(3.6)	629	(4.4)
	Saudi Arabia	399	(3.0)	84	(1.6)	256	(4.8)	286	(4.4)	341	(4.0)	402	(3.4)	459	(3.1)	507	(3.0)	534	(3.5)
	Serbia	439	(3.3)	96	(1.4)	282	(4.0)	312	(3.9)	370	(4.4)	440	(4.1)	508	(3.5)	566	(3.5)	599	(3.8)
	Singapore	549	(1.6)	109	(1.0)	352	(3.8)	398	(3.9)	478	(2.3)	559	(2.1)	628	(2.0)	684	(2.5)	714	(2.6)
	Chinese Taipei	503	(2.8)	102	(1.5)	325	(4.2)	367	(3.8)	435	(3.4)	508	(3.1)	576	(3.7)	630	(3.8)	661	(4.5)
	Thailand	393	(3.2)	79	(1.6)	271	(3.4)	295	(3.2)	337	(3.2)	388	(3.5)	445	(4.4)	501	(5.1)	533	(5.8)
	Ukraine	466	(3.5)	93	(1.7)	302	(6.2)	340	(5.2)	404	(4.8)	472	(3.5)	532	(3.5)	582	(3.8)	612	(4.8)
	United Arab Emirates	432	(2.3)	113	(0.9)	251	(2.4)	284	(2.7)	348	(2.5)	429	(2.6)	511	(3.5)	584	(3.1)	624	(3.0)
	Uruguay	427	(2.8)	96	(1.6)	267	(3.5)	299	(3.6)	360	(3.6)	427	(3.2)	495	(3.6)	552	(4.5)	585	(4.1)
	Viet Nam**	505	(3.6)	74	(1.7)	381	(4.9)	409	(4.6)	456	(3.9)	505	(3.7)	554	(4.1)	599	(4.7)	625	(5.5)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.5 [1/2] **Mean score and variation in mathematics performance**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD	Australia	491	(1.9)	92	(1.2)	339	(3.8)	371	(3.0)	428	(2.2)	492	(2.1)	555	(2.0)	609	(2.7)	641	(3.6)
	Austria	499	(3.0)	93	(1.5)	341	(4.4)	374	(4.4)	433	(4.0)	503	(3.7)	566	(3.5)	618	(3.3)	646	(3.6)
	Belgium	508	(2.3)	95	(1.7)	344	(4.3)	377	(4.1)	440	(3.2)	514	(2.5)	579	(2.6)	628	(3.4)	656	(3.7)
	Canada	512	(2.4)	92	(1.1)	358	(3.2)	392	(3.0)	449	(2.8)	513	(2.6)	576	(2.7)	629	(2.7)	661	(3.2)
	Chile	417	(2.4)	85	(1.4)	282	(3.9)	311	(3.5)	359	(2.9)	416	(2.9)	475	(3.2)	528	(3.5)	559	(4.1)
	Colombia	391	(3.0)	81	(2.0)	262	(5.4)	290	(3.9)	335	(3.5)	387	(3.5)	445	(3.8)	499	(4.5)	531	(4.4)
	Czech Republic	499	(2.5)	93	(1.7)	345	(5.2)	378	(4.6)	435	(3.6)	501	(2.7)	564	(2.8)	619	(3.1)	650	(3.9)
	Denmark	509	(1.7)	82	(1.0)	370	(3.6)	401	(2.6)	454	(2.3)	512	(2.3)	567	(2.3)	613	(2.8)	640	(3.5)
	Estonia	523	(1.7)	82	(1.1)	390	(3.1)	419	(2.9)	468	(2.4)	524	(2.0)	579	(2.2)	628	(2.7)	657	(3.6)
	Finland	507	(2.0)	82	(1.2)	368	(3.6)	399	(3.4)	451	(2.5)	510	(2.5)	565	(2.4)	612	(2.5)	639	(3.3)
	France	495	(2.3)	93	(1.5)	333	(4.3)	370	(3.4)	433	(3.2)	502	(3.0)	562	(3.2)	611	(3.3)	638	(3.6)
	Germany	500	(2.6)	95	(1.5)	337	(4.6)	373	(4.2)	433	(3.6)	504	(3.5)	570	(3.3)	621	(3.2)	650	(3.4)
	Greece	451	(3.1)	89	(1.8)	302	(4.9)	334	(4.7)	391	(4.1)	454	(3.3)	513	(3.2)	565	(3.8)	595	(4.7)
	Hungary	481	(2.3)	91	(1.6)	328	(3.9)	360	(4.0)	418	(3.3)	484	(2.9)	546	(3.0)	597	(3.7)	626	(4.7)
	Iceland	495	(2.0)	90	(1.2)	340	(3.8)	374	(4.2)	434	(3.4)	499	(2.7)	559	(2.7)	609	(3.0)	638	(4.1)
	Ireland	500	(2.2)	78	(1.0)	367	(3.6)	397	(3.3)	447	(2.6)	502	(2.5)	554	(2.3)	599	(3.0)	625	(3.5)
	Israel	463	(3.5)	108	(1.9)	276	(6.2)	315	(5.5)	388	(5.0)	468	(4.0)	542	(3.6)	600	(3.9)	632	(3.9)
	Italy	487	(2.8)	94	(1.8)	327	(5.5)	363	(4.7)	423	(3.1)	490	(3.5)	552	(3.3)	605	(3.9)	635	(4.9)
	Japan	527	(2.5)	86	(1.6)	380	(4.3)	413	(3.9)	468	(3.1)	530	(2.9)	589	(2.8)	637	(3.8)	664	(4.5)
	Korea	526	(3.1)	100	(2.0)	354	(5.0)	393	(4.4)	460	(3.8)	530	(3.4)	596	(3.6)	651	(4.6)	684	(5.9)
	Latvia	496	(2.0)	80	(1.1)	363	(4.1)	393	(3.2)	441	(2.4)	497	(2.4)	551	(2.5)	599	(3.1)	628	(3.4)
	Lithuania	481	(2.0)	91	(1.1)	330	(4.1)	362	(3.6)	418	(2.8)	483	(2.3)	545	(2.2)	598	(2.8)	630	(3.2)
	Luxembourg	483	(1.1)	98	(1.3)	321	(3.4)	353	(2.9)	413	(2.1)	485	(2.0)	555	(2.0)	611	(2.4)	641	(2.9)
	Mexico	409	(2.5)	78	(1.6)	284	(3.8)	311	(3.6)	356	(2.7)	408	(2.7)	461	(3.1)	510	(3.6)	539	(4.5)
	Netherlands*	519	(2.6)	93	(1.8)	362	(5.0)	394	(4.8)	453	(4.0)	524	(3.0)	588	(2.7)	638	(3.6)	664	(3.7)
	New Zealand	494	(1.7)	93	(1.1)	339	(3.7)	372	(3.0)	430	(2.5)	496	(2.3)	560	(2.2)	614	(2.2)	645	(3.7)
	Norway	501	(2.2)	90	(1.3)	345	(4.1)	381	(3.9)	441	(2.9)	504	(2.8)	565	(2.4)	617	(3.1)	645	(4.4)
	Poland	516	(2.6)	90	(1.7)	366	(4.7)	398	(3.8)	455	(2.9)	517	(2.8)	578	(3.1)	631	(4.2)	661	(4.7)
	Portugal*	492	(2.7)	96	(1.3)	327	(5.2)	362	(3.8)	426	(3.6)	497	(3.2)	562	(3.0)	614	(3.6)	643	(4.5)
	Slovak Republic	486	(2.6)	100	(1.7)	315	(6.0)	353	(5.4)	420	(4.1)	492	(3.0)	556	(2.7)	610	(3.1)	640	(3.7)
	Slovenia	509	(1.4)	89	(1.4)	360	(5.3)	392	(3.0)	448	(2.3)	511	(1.8)	571	(2.3)	622	(2.8)	652	(3.4)
	Spain	481	(1.5)	88	(1.0)	331	(2.8)	365	(2.4)	421	(1.8)	484	(1.6)	544	(1.8)	593	(2.2)	621	(2.4)
	Sweden	502	(2.7)	91	(1.4)	348	(5.7)	383	(4.6)	441	(3.7)	505	(3.2)	567	(2.9)	618	(3.3)	647	(3.8)
	Switzerland	515	(2.9)	94	(1.4)	360	(4.4)	391	(3.5)	448	(3.8)	518	(3.7)	582	(3.4)	636	(4.3)	668	(4.8)
	Turkey	454	(2.3)	88	(1.8)	314	(4.3)	343	(3.8)	392	(3.2)	450	(2.4)	512	(2.7)	571	(4.0)	605	(5.3)
	United Kingdom	502	(2.6)	93	(1.4)	346	(4.1)	381	(4.0)	439	(2.9)	504	(2.7)	567	(3.0)	620	(3.3)	651	(4.2)
	United States*	478	(3.2)	92	(1.5)	326	(5.0)	357	(4.6)	414	(4.0)	479	(3.8)	543	(3.9)	598	(4.3)	629	(4.6)
	OECD average	489	(0.4)	91	(0.2)	337	(0.7)	370	(0.6)	427	(0.5)	492	(0.5)	553	(0.5)	605	(0.6)	634	(0.7)
	OECD total	478	(1.0)	97	(0.5)	318	(1.7)	350	(1.5)	409	(1.3)	478	(1.2)	547	(1.2)	604	(1.3)	636	(1.5)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.5 [2/2] **Mean score and variation in mathematics performance**

	Mean score		Standard deviation		Percentiles													
					5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners																		
Albania	437	(2.4)	83	(1.3)	303	(3.6)	332	(3.1)	381	(2.9)	436	(3.0)	493	(2.8)	544	(3.5)	575	(3.8)
Argentina	379	(2.8)	84	(1.7)	243	(4.6)	272	(4.1)	322	(3.6)	378	(3.1)	436	(3.5)	489	(3.8)	520	(4.0)
Baku (Azerbaijan)	420	(2.8)	89	(1.7)	276	(3.8)	306	(3.4)	359	(2.9)	418	(3.1)	480	(3.8)	535	(5.0)	570	(5.4)
Belarus	472	(2.7)	93	(1.4)	318	(5.0)	351	(3.4)	407	(3.1)	473	(3.0)	537	(3.2)	592	(3.5)	623	(4.1)
Bosnia and Herzegovina	406	(3.1)	82	(1.3)	276	(4.1)	303	(3.2)	349	(3.2)	404	(3.5)	462	(3.7)	514	(4.4)	545	(4.3)
Brazil	384	(2.0)	88	(1.6)	251	(3.1)	277	(2.5)	322	(2.3)	377	(2.4)	440	(2.8)	501	(3.9)	538	(4.9)
Brunei Darussalam	430	(1.2)	91	(1.0)	287	(3.4)	316	(2.4)	365	(2.0)	425	(1.6)	492	(2.0)	555	(2.2)	588	(3.4)
B-S-J-Z (China)	591	(2.5)	80	(1.8)	452	(5.2)	486	(4.2)	540	(3.0)	596	(2.7)	647	(3.0)	691	(3.2)	716	(3.6)
Bulgaria	436	(3.8)	97	(2.1)	280	(6.1)	311	(4.6)	368	(4.6)	434	(4.2)	503	(4.1)	563	(5.7)	599	(6.8)
Costa Rica	402	(3.3)	75	(2.0)	282	(4.2)	308	(3.4)	352	(2.7)	401	(3.3)	452	(4.2)	499	(5.5)	528	(7.0)
Croatia	464	(2.5)	87	(1.7)	323	(4.6)	354	(3.9)	405	(3.0)	463	(2.9)	523	(3.1)	577	(3.9)	608	(4.2)
Cyprus	451	(1.4)	95	(1.1)	292	(3.5)	325	(2.8)	385	(2.5)	454	(1.9)	517	(2.1)	571	(2.4)	601	(3.4)
Dominican Republic	325	(2.6)	71	(2.0)	214	(3.2)	236	(2.7)	276	(2.7)	322	(2.9)	370	(3.2)	417	(4.8)	449	(6.6)
Georgia	398	(2.6)	88	(1.6)	257	(3.9)	286	(3.5)	336	(2.9)	394	(2.8)	457	(3.7)	515	(4.4)	548	(6.0)
Hong Kong (China)*	551	(3.0)	94	(1.9)	387	(6.2)	426	(5.4)	490	(4.2)	557	(3.1)	617	(2.8)	667	(3.5)	696	(4.5)
Indonesia	379	(3.1)	79	(2.2)	255	(4.3)	281	(3.9)	325	(3.2)	376	(3.1)	427	(3.7)	480	(5.9)	517	(8.7)
Jordan	400	(3.3)	85	(1.7)	259	(4.6)	291	(4.2)	343	(3.4)	400	(3.4)	458	(3.9)	508	(4.3)	539	(5.2)
Kazakhstan	423	(1.9)	87	(1.1)	282	(3.2)	314	(2.4)	365	(2.2)	422	(2.0)	480	(2.2)	535	(3.0)	568	(3.1)
Kosovo	366	(1.5)	77	(1.3)	243	(3.7)	269	(2.7)	313	(2.1)	364	(1.8)	416	(2.3)	465	(3.3)	497	(4.0)
Lebanon	393	(4.0)	106	(1.6)	224	(5.2)	256	(4.8)	317	(5.1)	391	(5.0)	469	(5.0)	533	(4.7)	569	(4.7)
Macao (China)	558	(1.5)	81	(1.5)	420	(4.1)	452	(3.6)	505	(2.3)	561	(2.3)	613	(2.2)	659	(2.6)	685	(3.4)
Malaysia	440	(2.9)	83	(1.7)	307	(3.6)	335	(3.0)	383	(3.1)	438	(3.0)	496	(3.9)	550	(4.8)	580	(5.9)
Malta	472	(1.9)	102	(1.4)	297	(4.4)	334	(3.4)	401	(3.6)	478	(2.7)	545	(2.7)	599	(3.5)	630	(4.8)
Moldova	421	(2.4)	94	(1.7)	268	(3.8)	300	(3.1)	354	(2.6)	419	(2.7)	486	(3.2)	543	(4.4)	578	(5.7)
Montenegro	430	(1.2)	83	(1.0)	295	(2.8)	324	(2.2)	371	(1.9)	429	(1.7)	487	(1.6)	538	(2.1)	569	(3.1)
Morocco	368	(3.3)	76	(1.5)	249	(3.5)	273	(3.2)	314	(3.3)	363	(3.6)	418	(4.4)	469	(4.4)	499	(5.0)
North Macedonia	394	(1.6)	93	(1.2)	243	(3.9)	275	(2.9)	330	(2.1)	394	(2.4)	458	(2.2)	516	(3.5)	550	(4.4)
Panama	353	(2.7)	77	(2.1)	228	(5.0)	255	(3.9)	300	(2.9)	351	(2.7)	403	(3.6)	454	(5.5)	485	(6.3)
Peru	400	(2.6)	84	(1.5)	266	(3.4)	293	(3.1)	341	(2.9)	397	(2.9)	456	(3.5)	511	(4.1)	544	(5.1)
Philippines	353	(3.5)	78	(2.0)	229	(4.2)	255	(3.7)	299	(3.2)	349	(3.4)	403	(4.5)	456	(6.0)	488	(7.4)
Qatar	414	(1.2)	98	(0.9)	259	(2.8)	290	(2.2)	345	(1.6)	411	(1.8)	481	(1.6)	544	(2.1)	582	(2.5)
Romania	430	(4.9)	94	(2.1)	277	(5.7)	310	(5.4)	365	(4.7)	428	(5.7)	495	(6.1)	554	(6.9)	588	(7.2)
Russia	488	(3.0)	86	(1.9)	344	(5.5)	376	(4.3)	430	(4.0)	489	(3.1)	547	(3.3)	597	(3.9)	627	(4.2)
Saudi Arabia	373	(3.0)	79	(1.6)	246	(4.6)	273	(4.3)	319	(3.4)	372	(3.3)	426	(3.6)	475	(3.6)	505	(4.1)
Serbia	448	(3.2)	97	(1.7)	293	(5.3)	324	(4.3)	380	(3.9)	446	(3.8)	516	(3.8)	576	(3.9)	609	(3.9)
Singapore	569	(1.6)	94	(1.2)	401	(3.4)	441	(2.9)	508	(2.4)	576	(2.0)	636	(2.1)	684	(2.7)	713	(3.0)
Chinese Taipei	531	(2.9)	100	(1.7)	358	(4.6)	397	(3.9)	466	(3.8)	537	(3.1)	601	(3.5)	656	(4.4)	686	(5.3)
Thailand	419	(3.4)	88	(1.8)	282	(4.8)	310	(3.6)	358	(3.3)	414	(3.7)	475	(4.3)	535	(5.8)	572	(6.1)
Ukraine	453	(3.6)	94	(1.9)	297	(5.2)	331	(4.4)	390	(4.2)	454	(4.1)	517	(4.1)	573	(5.0)	607	(5.7)
United Arab Emirates	435	(2.1)	106	(1.2)	265	(3.9)	299	(3.2)	360	(2.8)	433	(2.7)	509	(2.6)	574	(2.4)	611	(3.2)
Uruguay	418	(2.6)	85	(1.7)	276	(4.4)	307	(3.5)	359	(3.1)	419	(3.4)	477	(3.7)	529	(3.9)	558	(4.4)
Viet Nam**	496	(4.0)	75	(2.0)	373	(5.0)	400	(5.0)	445	(4.4)	496	(4.1)	546	(4.5)	591	(5.1)	619	(5.8)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.6 [1/2] **Mean score and variation in science performance**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD	Australia	503	(1.8)	101	(1.1)	334	(2.7)	369	(2.6)	432	(2.2)	506	(2.3)	575	(2.2)	631	(2.7)	664	(3.8)
	Austria	490	(2.8)	96	(1.2)	332	(3.8)	361	(3.1)	420	(3.6)	493	(3.5)	560	(3.1)	614	(3.3)	642	(3.7)
	Belgium	499	(2.2)	99	(1.3)	328	(4.2)	363	(4.0)	428	(3.4)	505	(2.6)	571	(2.5)	624	(2.3)	652	(2.8)
	Canada	518	(2.2)	96	(1.0)	357	(2.6)	393	(2.3)	453	(2.5)	520	(2.6)	586	(2.6)	640	(2.5)	671	(3.6)
	Chile	444	(2.4)	83	(1.4)	309	(3.6)	336	(3.1)	385	(3.0)	442	(2.9)	502	(3.3)	553	(3.3)	584	(3.8)
	Colombia	413	(3.1)	82	(1.4)	287	(3.8)	311	(3.7)	355	(3.6)	409	(3.6)	469	(4.0)	524	(4.1)	555	(4.2)
	Czech Republic	497	(2.5)	94	(1.6)	341	(4.8)	373	(4.0)	430	(3.7)	497	(3.1)	564	(3.1)	620	(2.9)	651	(3.6)
	Denmark	493	(1.9)	91	(1.3)	337	(3.8)	372	(3.4)	431	(2.6)	496	(2.5)	558	(2.6)	609	(3.1)	637	(3.6)
	Estonia	530	(1.9)	88	(1.2)	384	(3.9)	417	(3.5)	469	(2.9)	531	(2.4)	591	(2.4)	644	(2.7)	674	(3.0)
	Finland	522	(2.5)	96	(1.3)	356	(4.4)	393	(4.1)	458	(3.2)	526	(2.9)	590	(2.8)	643	(2.9)	673	(3.8)
	France	493	(2.2)	96	(1.4)	330	(4.2)	364	(3.5)	425	(3.1)	497	(3.1)	563	(2.9)	615	(3.2)	644	(3.8)
	Germany	503	(2.9)	103	(1.6)	328	(5.2)	363	(4.0)	430	(3.9)	508	(3.9)	577	(3.5)	633	(3.3)	665	(3.3)
	Greece	452	(3.1)	86	(1.6)	309	(5.2)	338	(4.6)	392	(4.1)	453	(3.6)	513	(3.3)	561	(3.4)	591	(4.2)
	Hungary	481	(2.3)	94	(1.4)	325	(4.4)	356	(3.9)	412	(3.1)	484	(3.1)	549	(3.3)	602	(3.6)	631	(4.1)
	Iceland	475	(1.8)	91	(1.0)	325	(3.6)	354	(3.1)	410	(3.0)	476	(2.6)	540	(2.7)	594	(3.1)	623	(3.7)
	Ireland	496	(2.2)	88	(1.2)	348	(4.1)	380	(3.5)	435	(2.6)	498	(2.6)	558	(2.6)	610	(3.2)	639	(4.2)
	Israel	462	(3.6)	111	(1.9)	279	(5.6)	314	(5.0)	381	(5.1)	464	(5.0)	544	(3.7)	607	(3.8)	640	(4.0)
	Italy	468	(2.4)	90	(1.7)	316	(4.7)	348	(3.9)	407	(3.1)	470	(3.0)	532	(3.0)	583	(3.7)	612	(4.7)
	Japan	529	(2.6)	92	(1.6)	371	(4.5)	405	(4.4)	466	(3.7)	534	(2.9)	595	(3.0)	646	(3.5)	673	(3.9)
	Korea	519	(2.8)	98	(1.7)	352	(4.9)	388	(4.1)	453	(3.7)	524	(3.3)	589	(3.1)	642	(3.8)	672	(4.4)
	Latvia	487	(1.8)	84	(1.2)	347	(3.8)	377	(3.3)	429	(2.8)	489	(2.2)	546	(2.3)	595	(2.7)	623	(3.3)
	Lithuania	482	(1.6)	90	(1.0)	334	(3.6)	364	(2.9)	418	(2.8)	483	(2.2)	546	(1.8)	599	(2.3)	629	(3.0)
	Luxembourg	477	(1.2)	98	(1.2)	317	(3.6)	347	(2.6)	404	(2.1)	477	(1.7)	549	(2.2)	606	(2.9)	637	(3.8)
	Mexico	419	(2.6)	74	(1.6)	303	(4.3)	326	(3.9)	367	(2.7)	416	(2.7)	469	(3.0)	518	(4.3)	548	(4.5)
	Netherlands*	503	(2.8)	104	(1.9)	329	(5.5)	364	(5.2)	428	(4.5)	508	(3.7)	581	(3.1)	636	(3.5)	666	(3.8)
	New Zealand	508	(2.1)	102	(1.4)	336	(4.5)	371	(3.7)	437	(2.8)	512	(2.7)	582	(2.7)	640	(2.9)	670	(3.3)
	Norway	490	(2.3)	98	(1.2)	321	(4.5)	357	(3.9)	424	(3.3)	495	(2.5)	560	(2.8)	616	(2.9)	645	(3.4)
	Poland	511	(2.6)	92	(1.4)	359	(4.2)	392	(3.4)	448	(2.8)	511	(3.0)	576	(3.4)	630	(4.0)	660	(4.4)
	Portugal*	492	(2.8)	92	(1.3)	336	(5.6)	368	(4.3)	427	(3.6)	494	(3.0)	558	(3.1)	609	(3.5)	638	(4.1)
	Slovak Republic	464	(2.3)	96	(1.5)	307	(3.9)	338	(3.5)	397	(3.2)	464	(2.9)	531	(2.9)	589	(3.5)	622	(3.7)
	Slovenia	507	(1.3)	88	(1.1)	359	(3.3)	390	(3.4)	447	(2.1)	510	(1.9)	569	(1.9)	621	(2.8)	648	(3.7)
	Spain	483	(1.6)	89	(0.8)	334	(2.3)	365	(2.4)	421	(1.9)	485	(1.7)	547	(1.8)	598	(2.2)	627	(2.2)
	Sweden	499	(3.1)	98	(1.5)	333	(6.0)	368	(5.1)	431	(4.0)	503	(3.4)	570	(3.1)	624	(3.3)	655	(3.8)
	Switzerland	495	(3.0)	97	(1.4)	335	(3.9)	367	(3.5)	426	(3.8)	497	(3.8)	565	(4.0)	622	(4.6)	651	(4.0)
	Turkey	468	(2.0)	84	(1.6)	335	(3.4)	361	(3.1)	409	(2.8)	466	(2.3)	526	(2.4)	579	(3.9)	608	(4.8)
	United Kingdom	505	(2.6)	99	(1.4)	340	(4.7)	374	(3.8)	437	(3.2)	507	(2.7)	575	(3.2)	632	(3.2)	664	(3.7)
	United States*	502	(3.3)	99	(1.6)	336	(6.1)	371	(4.9)	433	(4.4)	505	(3.9)	574	(3.8)	629	(3.9)	660	(3.8)
	OECD average	489	(0.4)	94	(0.2)	333	(0.7)	365	(0.6)	423	(0.5)	491	(0.5)	555	(0.5)	609	(0.5)	639	(0.6)
	OECD total	486	(1.1)	99	(0.5)	325	(1.5)	357	(1.5)	415	(1.3)	486	(1.2)	558	(1.3)	616	(1.4)	648	(1.6)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.6 [2/2] **Mean score and variation in science performance**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners	Albania	417	(2.0)	74	(1.1)	298	(3.2)	323	(3.1)	366	(2.4)	416	(2.5)	466	(2.6)	514	(3.2)	541	(3.6)
	Argentina	404	(2.9)	90	(1.6)	261	(4.7)	291	(4.0)	340	(3.4)	401	(3.3)	466	(3.7)	523	(4.0)	555	(3.7)
	Baku (Azerbaijan)	398	(2.4)	74	(1.6)	281	(3.0)	305	(2.5)	347	(2.3)	395	(2.2)	446	(3.0)	494	(4.6)	524	(6.2)
	Belarus	471	(2.4)	85	(1.3)	331	(3.7)	361	(3.5)	412	(3.4)	472	(2.9)	531	(2.7)	581	(2.7)	610	(3.7)
	Bosnia and Herzegovina	398	(2.7)	77	(1.3)	278	(3.6)	302	(3.1)	344	(2.7)	396	(3.2)	451	(3.6)	499	(3.8)	528	(4.1)
	Brazil	404	(2.1)	90	(1.5)	268	(3.0)	292	(2.3)	338	(2.1)	396	(2.3)	464	(3.1)	527	(3.6)	563	(4.8)
	Brunei Darussalam	431	(1.2)	96	(1.1)	290	(2.6)	315	(2.0)	359	(1.9)	421	(1.5)	497	(1.7)	566	(2.8)	603	(2.8)
	B-S-J-Z (China)	590	(2.7)	83	(1.7)	448	(5.0)	482	(4.0)	536	(3.4)	594	(2.8)	649	(3.1)	695	(3.7)	721	(3.9)
	Bulgaria	424	(3.6)	95	(2.0)	279	(5.1)	305	(4.3)	355	(4.0)	418	(4.1)	490	(4.8)	552	(5.3)	587	(6.1)
	Costa Rica	416	(3.3)	73	(1.9)	300	(3.9)	324	(3.2)	364	(3.0)	414	(3.4)	466	(4.3)	512	(5.6)	540	(6.6)
	Croatia	472	(2.8)	90	(1.6)	327	(4.2)	356	(4.0)	409	(3.5)	471	(3.2)	536	(3.1)	590	(3.5)	622	(3.9)
	Cyprus	439	(1.4)	93	(1.1)	291	(3.3)	319	(2.6)	372	(2.7)	437	(2.2)	505	(2.2)	562	(2.2)	592	(2.9)
	Dominican Republic	336	(2.5)	71	(1.6)	231	(2.7)	250	(2.8)	286	(2.4)	329	(2.9)	379	(3.5)	431	(4.8)	463	(5.7)
	Georgia	383	(2.3)	81	(1.3)	255	(3.6)	281	(2.7)	326	(2.7)	379	(2.9)	437	(3.0)	491	(3.9)	522	(4.9)
	Hong Kong (China)*	517	(2.5)	86	(1.2)	364	(4.6)	401	(4.3)	461	(3.2)	522	(2.7)	577	(2.5)	623	(3.3)	650	(4.0)
	Indonesia	396	(2.4)	69	(1.7)	289	(3.2)	312	(3.0)	348	(2.6)	392	(2.6)	440	(3.1)	488	(4.6)	517	(5.7)
	Jordan	429	(2.9)	88	(1.5)	282	(5.5)	316	(4.4)	370	(3.7)	431	(3.0)	490	(3.1)	541	(3.4)	570	(3.9)
	Kazakhstan	397	(1.7)	76	(1.4)	284	(2.6)	307	(2.1)	346	(1.9)	391	(1.8)	442	(2.4)	498	(3.4)	533	(4.8)
	Kosovo	365	(1.2)	65	(0.8)	265	(2.6)	285	(2.5)	320	(1.5)	361	(1.6)	406	(1.7)	450	(2.6)	478	(3.8)
	Lebanon	384	(3.5)	95	(1.6)	237	(4.0)	265	(3.6)	315	(3.7)	377	(4.3)	449	(4.8)	513	(4.9)	549	(4.9)
	Macao (China)	544	(1.5)	83	(1.0)	402	(4.3)	434	(3.0)	489	(2.6)	547	(1.8)	601	(1.9)	648	(2.2)	674	(3.5)
	Malaysia	438	(2.7)	77	(1.5)	313	(3.6)	339	(2.9)	384	(2.7)	436	(2.8)	490	(3.4)	538	(4.3)	565	(5.2)
	Malta	457	(1.9)	107	(1.2)	278	(4.8)	314	(3.5)	380	(2.9)	460	(2.5)	534	(2.9)	594	(3.3)	628	(4.2)
	Moldova	428	(2.3)	89	(1.4)	285	(3.8)	314	(2.9)	365	(2.5)	427	(2.7)	492	(3.2)	546	(3.7)	575	(4.1)
	Montenegro	415	(1.3)	81	(1.0)	285	(2.7)	311	(2.2)	358	(1.6)	413	(1.5)	470	(2.0)	523	(2.2)	554	(3.0)
	Morocco	377	(3.0)	67	(1.2)	275	(2.9)	293	(2.7)	328	(2.8)	372	(3.7)	422	(4.0)	468	(3.9)	493	(3.8)
	North Macedonia	413	(1.4)	92	(1.2)	265	(3.2)	296	(2.5)	349	(2.0)	411	(2.0)	476	(2.4)	533	(3.1)	566	(3.9)
	Panama	365	(2.9)	85	(1.9)	230	(4.8)	259	(3.8)	305	(3.2)	361	(2.8)	420	(4.1)	478	(5.7)	514	(6.1)
	Peru	404	(2.7)	80	(1.5)	280	(3.9)	304	(3.0)	347	(2.6)	400	(2.9)	458	(3.6)	511	(4.4)	543	(5.3)
	Philippines	357	(3.2)	75	(2.3)	250	(3.3)	269	(3.1)	304	(2.6)	347	(3.1)	401	(4.5)	461	(6.6)	500	(8.3)
	Qatar	419	(0.9)	103	(0.9)	259	(2.6)	290	(1.5)	345	(1.4)	414	(1.4)	490	(1.5)	557	(2.1)	596	(2.7)
	Romania	426	(4.6)	90	(1.8)	282	(5.5)	312	(4.7)	362	(4.6)	424	(5.4)	488	(5.5)	545	(5.8)	577	(6.2)
	Russia	478	(2.9)	84	(1.7)	339	(4.7)	369	(4.1)	420	(3.6)	478	(3.2)	536	(3.2)	586	(3.7)	616	(4.0)
	Saudi Arabia	386	(2.8)	79	(1.4)	261	(4.4)	287	(3.2)	331	(3.3)	384	(3.2)	440	(3.4)	489	(3.6)	519	(4.3)
	Serbia	440	(3.0)	92	(1.3)	293	(3.8)	322	(3.9)	375	(3.8)	438	(3.9)	504	(3.6)	562	(4.0)	593	(3.7)
	Singapore	551	(1.5)	97	(1.0)	376	(3.5)	416	(3.2)	487	(2.7)	560	(2.1)	621	(1.6)	670	(1.8)	698	(2.7)
	Chinese Taipei	516	(2.9)	99	(1.5)	346	(4.3)	382	(3.9)	449	(3.7)	521	(3.2)	587	(3.7)	641	(4.0)	670	(4.1)
	Thailand	426	(3.2)	82	(1.6)	299	(3.7)	324	(3.2)	367	(3.0)	421	(3.5)	481	(4.4)	535	(5.2)	567	(5.8)
	Ukraine	469	(3.3)	91	(1.8)	319	(5.0)	351	(4.4)	406	(3.8)	469	(3.8)	532	(3.7)	588	(4.5)	619	(5.5)
	United Arab Emirates	434	(2.0)	103	(0.8)	272	(2.4)	302	(2.1)	358	(2.2)	430	(2.6)	506	(2.8)	572	(3.0)	609	(2.8)
	Uruguay	426	(2.5)	87	(1.4)	287	(3.2)	314	(3.1)	364	(2.9)	423	(3.2)	486	(3.6)	540	(3.9)	573	(4.0)
	Viet Nam**	543	(3.3)	77	(1.7)	418	(4.5)	445	(3.9)	492	(3.9)	543	(3.5)	595	(3.9)	642	(4.0)	669	(5.0)

*Data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

StatLink https://doi.org/10.1787/888934029090

Table I.B1.7 [1/4] **Percentage of low achievers and top performers in reading, 2009 through 2018**

		Proficiency level in PISA 2009				Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	Australia	14.2	(0.6)	12.8	(0.8)	14.2	(0.5)	11.7	(0.5)	18.1	(0.5)	11.0	(0.5)	19.6	(0.5)	13.0	(0.5)
	Austria	m	m	m	m	19.5	(1.1)	5.5	(0.6)	22.5	(1.0)	7.2	(0.6)	23.6	(1.0)	7.4	(0.5)
	Belgium	17.7	(0.9)	11.2	(0.6)	16.1	(0.8)	11.8	(0.6)	19.5	(0.9)	9.3	(0.6)	21.3	(0.9)	9.5	(0.5)
	Canada	10.3	(0.5)	12.8	(0.5)	10.9	(0.5)	12.9	(0.6)	10.7	(0.6)	14.0	(0.7)	13.8	(0.5)	15.0	(0.6)
	Chile	30.6	(1.5)	1.3	(0.3)	33.0	(1.7)	0.6	(0.1)	28.4	(1.2)	2.3	(0.3)	31.7	(1.2)	2.6	(0.3)
	Colombia	47.1	(1.9)	0.6	(0.2)	51.4	(1.8)	0.3	(0.1)	42.8	(1.5)	1.0	(0.2)	49.9	(1.7)	0.9	(0.2)
	Czech Republic	23.1	(1.3)	5.1	(0.5)	16.9	(1.2)	6.1	(0.5)	22.0	(1.1)	7.9	(0.6)	20.7	(1.1)	8.2	(0.5)
	Denmark	15.2	(0.9)	4.7	(0.5)	14.6	(1.1)	5.4	(0.6)	15.0	(0.8)	6.5	(0.6)	16.0	(0.7)	8.4	(0.5)
	Estonia	13.3	(1.0)	6.1	(0.6)	9.1	(0.6)	8.3	(0.7)	10.6	(0.7)	11.0	(0.7)	11.1	(0.6)	13.9	(0.7)
	Finland	8.1	(0.5)	14.5	(0.8)	11.3	(0.7)	13.5	(0.6)	11.1	(0.8)	13.7	(0.7)	13.5	(0.7)	14.2	(0.7)
	France	19.8	(1.2)	9.6	(1.0)	18.9	(1.0)	12.9	(0.8)	21.5	(0.9)	12.5	(0.7)	20.9	(0.7)	9.2	(0.7)
	Germany	18.5	(1.1)	7.6	(0.6)	14.5	(0.9)	8.9	(0.7)	16.2	(0.9)	11.7	(0.7)	20.7	(1.1)	11.3	(0.7)
	Greece	21.3	(1.8)	5.6	(0.5)	22.6	(1.2)	5.1	(0.6)	27.3	(1.8)	4.0	(0.5)	30.5	(1.5)	3.7	(0.5)
	Hungary	17.6	(1.4)	6.1	(0.7)	19.7	(1.2)	5.6	(0.8)	27.5	(1.1)	4.3	(0.4)	25.3	(0.9)	5.7	(0.5)
	Iceland	16.8	(0.6)	8.5	(0.6)	21.0	(0.7)	5.8	(0.5)	22.1	(1.0)	6.6	(0.6)	26.4	(0.9)	7.1	(0.6)
	Ireland	17.2	(1.0)	7.0	(0.5)	9.6	(0.9)	11.4	(0.7)	10.2	(0.8)	10.7	(0.7)	11.8	(0.7)	12.1	(0.7)
	Israel	26.5	(1.2)	7.4	(0.6)	23.6	(1.6)	9.6	(0.8)	26.6	(1.3)	9.2	(0.7)	31.1	(1.3)	10.4	(0.7)
	Italy	21.0	(0.6)	5.8	(0.3)	19.5	(0.7)	6.7	(0.3)	21.0	(1.0)	5.7	(0.5)	23.3	(1.0)	5.3	(0.5)
	Japan	13.6	(1.1)	13.4	(0.9)	9.8	(0.9)	18.5	(1.3)	12.9	(1.0)	10.8	(0.9)	16.8	(1.0)	10.3	(0.7)
	Korea	5.8	(0.8)	12.9	(1.1)	7.6	(0.9)	14.1	(1.2)	13.7	(1.0)	12.7	(1.0)	15.1	(0.9)	13.1	(0.9)
	Latvia	17.6	(1.2)	2.9	(0.4)	17.0	(1.1)	4.2	(0.6)	17.7	(0.9)	4.3	(0.5)	22.4	(0.7)	4.8	(0.4)
	Lithuania	24.4	(1.2)	2.9	(0.4)	21.2	(1.2)	3.3	(0.4)	25.1	(0.9)	4.4	(0.5)	24.4	(0.8)	5.0	(0.4)
	Luxembourg	26.0	(0.6)	5.7	(0.5)	22.2	(0.7)	8.9	(0.4)	25.6	(0.6)	8.1	(0.4)	29.3	(0.6)	7.6	(0.5)
	Mexico	40.1	(1.0)	0.4	(0.1)	41.1	(0.9)	0.4	(0.1)	41.7	(1.3)	0.3	(0.1)	44.7	(1.3)	0.8	(0.2)
	Netherlands*	14.3	(1.5)	9.8	(1.1)	14.0	(1.2)	9.8	(0.8)	18.1	(1.0)	10.9	(0.6)	24.1	(1.0)	9.1	(0.6)
	New Zealand	14.3	(0.7)	15.7	(0.8)	16.3	(0.8)	14.0	(0.8)	17.3	(0.8)	13.6	(0.9)	19.0	(0.8)	13.1	(0.6)
	Norway	15.0	(0.8)	8.4	(0.9)	16.2	(1.0)	10.2	(0.7)	14.9	(0.8)	12.2	(0.7)	19.3	(0.8)	11.3	(0.6)
	Poland	15.0	(0.8)	7.2	(0.6)	10.6	(0.8)	10.0	(0.9)	14.4	(0.8)	8.2	(0.7)	14.7	(0.8)	12.2	(0.8)
	Portugal*	17.6	(1.2)	4.8	(0.5)	18.8	(1.4)	5.8	(0.6)	17.2	(0.9)	7.5	(0.6)	20.2	(0.9)	7.3	(0.6)
	Slovak Republic	22.2	(1.2)	4.5	(0.5)	28.2	(1.8)	4.4	(0.7)	32.1	(1.1)	3.5	(0.4)	31.4	(1.0)	4.6	(0.4)
	Slovenia	21.2	(0.6)	4.6	(0.5)	21.1	(0.7)	5.0	(0.4)	15.1	(0.6)	8.9	(0.7)	17.9	(0.7)	7.8	(0.5)
	Spain	19.6	(0.9)	3.3	(0.3)	18.3	(0.8)	5.5	(0.3)	16.2	(0.9)	5.5	(0.5)	m	m	m	m
	Sweden	17.4	(0.9)	9.0	(0.7)	22.7	(1.2)	7.9	(0.6)	18.4	(1.1)	10.0	(0.8)	18.4	(1.0)	13.3	(0.7)
	Switzerland	16.8	(0.9)	8.1	(0.7)	13.7	(0.8)	9.1	(0.7)	20.0	(1.1)	7.8	(0.6)	23.6	(1.1)	8.1	(0.7)
	Turkey	24.5	(1.4)	1.9	(0.4)	21.6	(1.4)	4.3	(0.9)	40.0	(2.0)	0.6	(0.2)	26.1	(1.0)	3.3	(0.5)
	United Kingdom	18.4	(0.8)	8.0	(0.5)	16.6	(1.3)	8.8	(0.7)	17.9	(0.9)	9.2	(0.6)	17.3	(0.9)	11.5	(0.8)
	United States*	17.6	(1.1)	9.9	(0.9)	16.6	(1.3)	7.9	(0.7)	19.0	(1.1)	9.6	(0.7)	19.3	(1.1)	13.5	(0.9)
	OECD average-35a	19.4	(0.2)	7.3	(0.1)	18.9	(0.2)	8.1	(0.1)	20.9	(0.2)	8.1	(0.1)	22.6	(0.2)	8.8	(0.1)
	OECD average-36a	m	m	m	m	18.9	(0.2)	8.0	(0.1)	21.0	(0.2)	8.1	(0.1)	22.6	(0.2)	8.7	(0.1)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.7 [2/4] **Percentage of low achievers and top performers in reading, 2009 through 2018**

		Proficiency level in PISA 2009				Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners	**Albania**	56.7	(1.9)	0.2	(0.1)	52.3	(1.3)	1.2	(0.2)	50.3	(1.9)	1.0	(0.2)	52.2	(1.1)	0.4	(0.1)
	Argentina	51.6	(1.9)	1.0	(0.2)	53.6	(1.7)	0.5	(0.1)	m	m	m	m	52.1	(1.3)	0.7	(0.2)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	60.4	(1.3)	0.1	(0.1)
	Belarus	m	m	m	m	m	m	m	m	m	m	m	m	23.4	(1.0)	3.9	(0.4)
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	53.7	(1.6)	0.2	(0.1)
	Brazil	49.6	(1.3)	1.3	(0.2)	50.8	(1.1)	0.5	(0.1)	51.0	(1.1)	1.4	(0.2)	50.0	(0.9)	1.8	(0.2)
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	51.8	(0.6)	1.3	(0.2)
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	5.2	(0.6)	21.7	(1.1)
	Bulgaria	41.0	(2.6)	2.8	(0.5)	39.4	(2.2)	4.3	(0.6)	41.5	(2.0)	3.6	(0.5)	47.1	(1.7)	2.3	(0.4)
	Costa Rica	32.6	(1.5)	0.8	(0.3)	32.4	(1.8)	0.6	(0.2)	40.3	(1.4)	0.7	(0.2)	42.0	(1.6)	0.6	(0.2)
	Croatia	22.4	(1.3)	3.2	(0.4)	18.7	(1.3)	4.4	(0.7)	19.9	(1.1)	5.9	(0.5)	21.6	(1.2)	4.7	(0.5)
	Cyprus	m	m	m	m	32.8	(0.7)	4.0	(0.3)	35.6	(0.9)	3.1	(0.3)	43.7	(0.7)	1.8	(0.2)
	Dominican Republic	m	m	m	m	m	m	m	m	72.1	(1.5)	0.1	(0.1)	79.1	(1.3)	0.1	(0.1)
	Georgia	62.0	(1.3)	0.3	(0.1)	m	m	m	m	51.7	(1.3)	1.1	(0.2)	64.4	(1.1)	0.2	(0.1)
	Hong Kong (China)*	8.3	(0.7)	12.4	(0.8)	6.8	(0.7)	16.8	(1.2)	9.3	(0.8)	11.6	(0.9)	12.6	(0.8)	14.8	(0.7)
	Indonesia	53.4	(2.3)	0.0	(0.0)	55.2	(2.2)	0.1	(0.1)	55.4	(1.5)	0.2	(0.1)	69.9	(1.4)	0.1	(0.0)
	Jordan	48.0	(1.6)	0.2	(0.1)	50.7	(1.6)	0.1	(0.1)	46.3	(1.4)	0.3	(0.1)	41.2	(1.4)	0.3	(0.1)
	Kazakhstan	58.7	(1.5)	0.4	(0.1)	57.1	(1.6)	0.0	(0.0)	m	m	m	m	64.2	(0.7)	0.4	(0.1)
	Kosovo	m	m	m	m	m	m	m	m	76.9	(0.9)	0.0	c	78.7	(0.6)	0.0	(0.0)
	Lebanon	m	m	m	m	m	m	m	m	70.4	(1.6)	0.8	(0.3)	67.8	(1.5)	0.7	(0.2)
	Macao (China)	14.9	(0.5)	2.9	(0.2)	11.5	(0.4)	7.0	(0.4)	11.7	(0.5)	6.7	(0.5)	10.8	(0.5)	13.8	(0.6)
	Malaysia	44.0	(1.6)	0.1	(0.1)	52.7	(1.7)	0.1	(0.1)	m	m	m	m	45.8	(1.4)	0.5	(0.2)
	Malta	36.3	(0.7)	4.4	(0.4)	m	m	m	m	35.6	(0.8)	5.6	(0.4)	35.9	(0.8)	5.3	(0.5)
	Moldova	57.2	(1.5)	0.1	(0.1)	m	m	m	m	45.8	(1.1)	1.2	(0.2)	43.0	(1.1)	1.0	(0.3)
	Montenegro	49.5	(1.0)	0.6	(0.2)	43.3	(0.7)	1.0	(0.2)	41.9	(0.7)	1.4	(0.3)	44.4	(0.7)	0.8	(0.2)
	Morocco	m	m	m	m	m	m	m	m	m	m	m	m	73.3	(1.6)	0.0	(0.0)
	North Macedonia	m	m	m	m	m	m	m	m	70.7	(0.7)	0.2	(0.1)	55.1	(0.7)	0.3	(0.2)
	Panama	65.3	(2.6)	0.5	(0.2)	m	m	m	m	m	m	m	m	64.3	(1.4)	0.2	(0.1)
	Peru	64.8	(1.7)	0.5	(0.2)	59.9	(2.0)	0.5	(0.2)	53.9	(1.5)	0.3	(0.1)	54.3	(1.3)	0.8	(0.2)
	Philippines	m	m	m	m	m	m	m	m	m	m	m	m	80.6	(1.4)	0.1	(0.0)
	Qatar	63.5	(0.5)	1.7	(0.2)	57.1	(0.4)	1.6	(0.1)	51.6	(0.5)	1.6	(0.2)	50.9	(0.4)	2.6	(0.2)
	Romania	40.4	(2.0)	0.7	(0.2)	37.3	(1.9)	1.6	(0.4)	38.7	(1.9)	2.0	(0.4)	40.8	(2.2)	1.4	(0.3)
	Russia	27.4	(1.3)	3.2	(0.5)	22.3	(1.3)	4.6	(0.6)	16.2	(1.2)	6.7	(0.6)	22.1	(1.2)	5.4	(0.5)
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	52.4	(1.5)	0.1	(0.1)
	Serbia	32.8	(1.3)	0.8	(0.2)	33.1	(1.7)	2.2	(0.4)	m	m	m	m	37.7	(1.5)	2.5	(0.3)
	Singapore	12.5	(0.5)	15.7	(0.5)	9.9	(0.4)	21.2	(0.6)	11.1	(0.5)	18.4	(0.7)	11.2	(0.5)	25.8	(0.7)
	Chinese Taipei	15.6	(0.9)	5.2	(0.8)	11.5	(0.9)	11.8	(0.8)	17.2	(0.8)	6.9	(0.8)	17.8	(0.8)	10.9	(0.8)
	Thailand	42.9	(1.5)	0.3	(0.2)	33.0	(1.4)	0.8	(0.2)	50.0	(1.8)	0.3	(0.1)	59.5	(1.7)	0.2	(0.1)
	Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	25.9	(1.4)	3.4	(0.5)
	United Arab Emirates	39.8	(1.2)	2.3	(0.3)	35.5	(1.1)	2.2	(0.3)	40.4	(1.2)	3.0	(0.3)	42.9	(0.8)	4.8	(0.3)
	Uruguay	41.9	(1.2)	1.8	(0.3)	47.0	(1.4)	0.9	(0.3)	39.0	(1.1)	2.5	(0.4)	41.9	(1.3)	1.5	(0.3)
	Viet Nam**	m	m	m	m	9.4	(1.4)	4.5	(0.8)	13.8	(1.4)	2.7	(0.7)	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.7 [3/4] **Percentage of low achievers and top performers in reading, 2009 through 2018**

	Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
	Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)	
	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
OECD												
Australia	**5.4**	(1.1)	0.3	(1.1)	**5.4**	(1.1)	1.3	(1.1)	1.6	(1.2)	2.0	(1.1)
Austria	m	m	m	m	**4.1**	(1.7)	**1.9**	(0.9)	1.1	(1.8)	0.2	(0.9)
Belgium	**3.5**	(1.6)	-1.6	(0.9)	**5.2**	(1.6)	**-2.2**	(1.0)	1.7	(1.7)	0.2	(1.0)
Canada	**3.5**	(0.9)	2.2	(1.2)	**2.9**	(0.9)	2.1	(1.4)	**3.1**	(1.1)	1.0	(1.6)
Chile	1.2	(2.7)	**1.3**	(0.4)	-1.3	(2.9)	**2.0**	(0.3)	3.3	(2.9)	0.3	(0.4)
Colombia	2.8	(3.4)	0.4	(0.2)	-1.5	(3.5)	**0.6**	(0.2)	**7.1**	(3.6)	0.0	(0.2)
Czech Republic	-2.3	(2.0)	**3.1**	(0.8)	3.9	(2.0)	**2.2**	(0.9)	-1.3	(2.0)	0.4	(0.9)
Denmark	0.8	(1.4)	**3.7**	(0.9)	1.4	(1.6)	**3.0**	(1.1)	1.0	(1.5)	1.9	(1.1)
Estonia	-2.3	(1.3)	**7.8**	(1.2)	1.9	(1.1)	**5.5**	(1.3)	0.4	(1.1)	**2.8**	(1.4)
Finland	**5.4**	(1.0)	-0.3	(1.5)	**2.2**	(1.1)	0.7	(1.6)	**2.5**	(1.1)	0.5	(1.7)
France	1.2	(1.6)	-0.4	(1.3)	2.0	(1.5)	**-3.7**	(1.3)	-0.5	(1.5)	**-3.3**	(1.2)
Germany	2.2	(1.7)	**3.7**	(1.1)	**6.2**	(1.6)	2.4	(1.2)	**4.5**	(1.7)	-0.4	(1.3)
Greece	**9.2**	(2.8)	**-2.0**	(0.7)	**7.9**	(2.6)	-1.5	(0.8)	3.2	(3.0)	-0.4	(0.7)
Hungary	**7.7**	(2.1)	-0.4	(0.9)	**5.6**	(2.1)	0.1	(1.0)	-2.2	(2.1)	1.4	(0.8)
Iceland	**9.5**	(1.6)	-1.4	(0.9)	**5.4**	(1.7)	1.3	(0.8)	**4.3**	(2.0)	0.5	(0.9)
Ireland	**-5.4**	(1.4)	**5.1**	(1.3)	2.2	(1.3)	0.7	(1.4)	1.6	(1.3)	1.4	(1.5)
Israel	**4.5**	(1.9)	**3.0**	(1.0)	**7.5**	(2.3)	0.8	(1.2)	**4.5**	(2.0)	1.2	(1.1)
Italy	2.2	(1.7)	-0.5	(0.7)	**3.8**	(1.9)	-1.3	(0.7)	2.3	(2.2)	-0.3	(0.8)
Japan	3.2	(1.7)	**-3.2**	(1.2)	**7.0**	(1.6)	**-8.2**	(1.6)	**3.9**	(1.7)	-0.5	(1.3)
Korea	**9.3**	(1.3)	0.2	(1.6)	**7.5**	(1.4)	-1.0	(1.7)	1.5	(1.5)	0.5	(1.6)
Latvia	**4.9**	(2.0)	**1.9**	(0.7)	**5.5**	(2.1)	0.7	(0.8)	**4.8**	(2.1)	0.5	(0.7)
Lithuania	0.0	(1.8)	**2.1**	(0.6)	3.2	(1.9)	**1.7**	(0.6)	-0.7	(1.8)	0.5	(0.7)
Luxembourg	**3.3**	(1.3)	**1.9**	(0.7)	**7.1**	(1.4)	-1.3	(0.7)	**3.6**	(1.5)	-0.5	(0.8)
Mexico	4.6	(3.6)	0.4	(0.2)	3.6	(3.9)	0.4	(0.2)	2.9	(4.3)	**0.5**	(0.2)
Netherlands*	**9.8**	(2.2)	-0.7	(1.3)	**10.1**	(2.1)	-0.7	(1.2)	**6.0**	(2.1)	-1.8	(1.1)
New Zealand	**4.6**	(1.4)	-2.6	(1.4)	2.7	(1.5)	-0.9	(1.4)	1.7	(1.6)	-0.5	(1.6)
Norway	**4.3**	(1.3)	**2.9**	(1.2)	**3.1**	(1.5)	1.1	(1.1)	**4.4**	(1.4)	-0.9	(1.1)
Poland	-0.4	(1.4)	**5.0**	(1.3)	**4.1**	(1.4)	2.2	(1.6)	0.3	(1.5)	**4.0**	(1.5)
Portugal*	2.6	(1.8)	**2.5**	(0.9)	1.4	(2.0)	1.5	(1.0)	3.0	(1.8)	-0.2	(1.0)
Slovak Republic	**9.2**	(2.3)	0.2	(0.7)	3.2	(2.8)	0.3	(0.8)	-0.7	(2.6)	1.2	(0.6)
Slovenia	**-3.3**	(1.2)	**3.2**	(0.8)	**-3.3**	(1.3)	**2.7**	(0.8)	**2.7**	(1.3)	-1.2	(0.9)
Spain	m	m	m	m	m	m	m	m	m	m	m	m
Sweden	1.0	(1.5)	**4.2**	(1.4)	**-4.3**	(1.7)	**5.4**	(1.5)	0.0	(1.7)	**3.3**	(1.7)
Switzerland	**6.8**	(1.6)	0.0	(1.1)	**9.9**	(1.7)	-1.0	(1.1)	3.7	(1.9)	0.3	(1.1)
Turkey	1.6	(2.6)	**1.5**	(0.6)	4.5	(2.8)	-1.0	(1.0)	**-13.8**	(3.3)	**2.7**	(0.6)
United Kingdom	-1.1	(1.5)	**3.4**	(1.0)	0.7	(1.9)	**2.7**	(1.2)	-0.6	(1.7)	2.3	(1.2)
United States*	1.6	(1.7)	**3.7**	(1.5)	2.7	(1.8)	**5.6**	(1.4)	0.3	(1.8)	**4.0**	(1.5)
OECD average-35a	**3.2**	(1.0)	**1.4**	(0.5)	**3.7**	(1.2)	0.7	(0.6)	1.7	(1.3)	0.7	(0.6)
OECD average-36a	m	m	m	m	**3.7**	(1.1)	0.7	(0.6)	1.7	(1.3)	0.6	(0.5)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.7 [4/4] **Percentage of low achievers and top performers in reading, 2009 through 2018**

	Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
	Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)		Below Level 2 (less than 407.47 score points)		Level 5 or above (at or above 625.61 score points)	
	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
Partners												
Albania	-4.4	(4.0)	0.2	(0.2)	-0.1	(4.2)	**-0.8**	(0.3)	2.0	(4.8)	**-0.6**	(0.3)
Argentina	0.5	(3.2)	-0.3	(0.3)	-1.4	(3.3)	0.2	(0.2)	m	m	m	m
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	0.4	(2.4)	0.5	(0.3)	-0.8	(2.5)	**1.3**	(0.3)	-1.0	(2.6)	0.4	(0.3)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	6.1	(3.5)	-0.4	(0.6)	**7.7**	(3.4)	**-2.0**	(0.7)	5.6	(3.3)	**-1.2**	(0.6)
Costa Rica	**9.3**	(3.8)	-0.2	(0.3)	**9.6**	(4.2)	0.1	(0.3)	1.7	(4.4)	0.0	(0.2)
Croatia	-0.9	(2.1)	**1.5**	(0.7)	2.9	(2.2)	0.3	(0.9)	1.7	(2.2)	-1.2	(0.8)
Cyprus	m	m	m	m	**10.9**	(2.2)	**-2.2**	(0.4)	**8.1**	(2.5)	**-1.3**	(0.4)
Dominican Republic	m	m	m	m	m	m	m	m	**6.9**	(2.4)	0.0	(0.1)
Georgia	2.4	(2.9)	-0.1	(0.1)	m	m	m	m	**12.7**	(3.4)	**-0.9**	(0.3)
Hong Kong (China)*	**4.3**	(1.1)	2.4	(1.4)	**5.8**	(1.1)	-2.0	(1.7)	**3.3**	(1.1)	**3.3**	(1.6)
Indonesia	**16.5**	(3.7)	0.0	(0.0)	**14.7**	(3.9)	0.0	(0.1)	**14.5**	(3.8)	-0.1	(0.1)
Jordan	-6.8	(3.8)	0.0	(0.1)	**-9.5**	(4.1)	0.1	(0.2)	-5.1	(4.3)	0.0	(0.1)
Kazakhstan	5.5	(3.5)	0.0	(0.1)	7.1	(3.9)	**0.4**	(0.1)	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	1.9	(2.9)	0.0	(0.0)
Lebanon	m	m	m	m	m	m	m	m	-2.6	(2.5)	-0.1	(0.3)
Macao (China)	**-4.1**	(0.9)	**10.9**	(1.0)	-0.7	(0.9)	**6.8**	(1.1)	-0.9	(0.9)	**7.1**	(1.3)
Malaysia	1.9	(3.7)	0.4	(0.2)	-6.9	(4.0)	0.4	(0.2)	m	m	m	m
Malta	-0.4	(1.6)	0.9	(0.6)	m	m	m	m	0.3	(1.9)	-0.3	(0.7)
Moldova	**-14.2**	(2.9)	**0.9**	(0.3)	m	m	m	m	-2.8	(3.2)	-0.2	(0.4)
Montenegro	-5.1	(2.9)	0.2	(0.3)	1.1	(3.2)	-0.2	(0.3)	2.5	(3.5)	-0.6	(0.4)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	m	m	**-15.5**	(2.9)	0.2	(0.2)
Panama	-0.9	(3.7)	-0.3	(0.2)	m	m	m	m	m	m	m	m
Peru	**-10.5**	(2.9)	0.3	(0.3)	-5.6	(3.2)	0.3	(0.3)	0.4	(3.2)	**0.5**	(0.2)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	**-12.6**	(1.7)	**0.9**	(0.2)	**-6.3**	(1.9)	**1.0**	(0.2)	-0.7	(2.1)	**1.1**	(0.2)
Romania	0.4	(3.5)	**0.7**	(0.4)	3.6	(3.5)	-0.2	(0.5)	2.1	(3.7)	-0.6	(0.5)
Russia	**-5.3**	(2.2)	**2.3**	(0.7)	-0.2	(2.3)	0.8	(0.8)	**5.9**	(2.3)	-1.2	(0.8)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	4.9	(2.7)	**1.7**	(0.4)	4.6	(3.0)	0.3	(0.5)	m	m	m	m
Singapore	-1.2	(0.7)	**10.1**	(1.8)	**1.4**	(0.7)	**4.6**	(2.0)	0.1	(0.8)	**7.5**	(2.2)
Chinese Taipei	2.2	(1.4)	**5.7**	(1.3)	**6.3**	(1.5)	-0.9	(1.4)	0.6	(1.5)	**4.0**	(1.4)
Thailand	**16.7**	(3.7)	-0.1	(0.2)	**26.6**	(3.9)	**-0.6**	(0.2)	**9.6**	(4.4)	-0.1	(0.1)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	3.1	(2.0)	**2.5**	(0.4)	**7.4**	(2.0)	**2.6**	(0.4)	2.5	(2.2)	**1.8**	(0.5)
Uruguay	0.0	(2.9)	-0.2	(0.4)	-5.2	(3.2)	0.6	(0.4)	2.9	(3.3)	**-1.0**	(0.5)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [1/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

		Proficiency level in PISA 2003				Proficiency level in PISA 2006				Proficiency level in PISA 2009			
		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	Australia	14.3	(0.7)	19.8	(0.8)	13.0	(0.6)	16.4	(0.8)	15.9	(0.7)	16.4	(0.9)
	Austria	18.8	(1.2)	14.3	(1.0)	20.0	(1.4)	15.8	(1.0)	m	m	m	m
	Belgium	16.5	(0.8)	26.4	(0.8)	17.3	(1.0)	22.3	(0.8)	19.1	(0.8)	20.4	(0.7)
	Canada	10.1	(0.5)	20.3	(0.7)	10.8	(0.6)	17.9	(0.7)	11.5	(0.5)	18.3	(0.6)
	Chile	m	m	m	m	55.1	(2.2)	1.5	(0.4)	51.0	(1.7)	1.3	(0.3)
	Colombia	m	m	m	m	71.9	(1.6)	0.4	(0.2)	70.4	(1.6)	0.1	(0.1)
	Czech Republic	16.6	(1.3)	18.3	(1.2)	19.2	(1.2)	18.3	(1.2)	22.3	(1.1)	11.6	(0.9)
	Denmark	15.4	(0.8)	15.9	(0.9)	13.6	(1.0)	13.7	(0.8)	17.1	(0.9)	11.6	(0.8)
	Estonia	m	m	m	m	12.1	(1.0)	12.5	(0.8)	12.6	(0.9)	12.1	(0.8)
	Finland	6.8	(0.5)	23.4	(0.8)	6.0	(0.6)	24.4	(1.0)	7.8	(0.5)	21.7	(0.9)
	France	16.6	(1.1)	15.1	(0.9)	22.3	(1.3)	12.5	(0.9)	22.5	(1.3)	13.7	(1.0)
	Germany	21.6	(1.2)	16.2	(0.9)	19.9	(1.4)	15.4	(1.0)	18.6	(1.1)	17.8	(0.9)
	Greece	38.9	(1.9)	4.0	(0.6)	32.3	(1.4)	5.0	(0.5)	30.3	(1.8)	5.7	(0.6)
	Hungary	23.0	(1.0)	10.7	(0.9)	21.2	(1.1)	10.3	(0.9)	22.3	(1.5)	10.1	(1.1)
	Iceland	15.0	(0.7)	15.5	(0.7)	16.8	(0.8)	12.7	(0.7)	17.0	(0.6)	13.6	(0.6)
	Ireland	16.8	(1.0)	11.4	(0.8)	16.4	(1.2)	10.2	(0.8)	20.8	(1.0)	6.7	(0.6)
	Israel	m	m	m	m	42.0	(1.7)	6.1	(0.6)	39.5	(1.3)	5.9	(0.7)
	Italy	31.9	(1.5)	7.0	(0.5)	32.8	(0.9)	6.2	(0.5)	24.9	(0.6)	9.0	(0.5)
	Japan	13.3	(1.2)	24.3	(1.5)	13.0	(1.1)	18.3	(1.0)	12.5	(1.0)	20.9	(1.2)
	Korea	9.5	(0.8)	24.8	(1.4)	8.9	(1.0)	27.1	(1.5)	8.1	(1.0)	25.6	(1.6)
	Latvia	23.7	(1.4)	8.0	(0.8)	20.7	(1.2)	6.6	(0.6)	22.6	(1.4)	5.7	(0.6)
	Lithuania	m	m	m	m	23.0	(1.1)	9.1	(0.9)	26.3	(1.2)	7.0	(0.7)
	Luxembourg	21.7	(0.6)	10.8	(0.6)	22.8	(0.6)	10.6	(0.5)	23.9	(0.6)	11.4	(0.6)
	Mexico	65.9	(1.7)	0.4	(0.1)	56.5	(1.3)	0.8	(0.2)	50.8	(1.0)	0.7	(0.1)
	Netherlands*	10.9	(1.1)	25.5	(1.3)	11.5	(1.0)	21.1	(1.1)	13.4	(1.4)	19.9	(1.5)
	New Zealand	15.1	(0.8)	20.7	(0.7)	14.0	(0.8)	18.9	(0.9)	15.4	(0.9)	18.9	(0.9)
	Norway	20.8	(1.0)	11.4	(0.6)	22.2	(1.2)	10.4	(0.7)	18.2	(0.9)	10.2	(0.7)
	Poland	22.0	(1.1)	10.1	(0.6)	19.8	(0.9)	10.6	(0.8)	20.5	(1.1)	10.4	(0.9)
	Portugal*	30.1	(1.7)	5.4	(0.5)	30.7	(1.5)	5.7	(0.5)	23.7	(1.1)	9.6	(0.8)
	Slovak Republic	19.9	(1.4)	12.7	(0.9)	20.9	(1.0)	11.0	(0.9)	21.0	(1.2)	12.7	(1.0)
	Slovenia	m	m	m	m	17.7	(0.7)	13.7	(0.6)	20.3	(0.5)	14.2	(0.6)
	Spain	23.0	(1.0)	7.9	(0.7)	24.7	(1.1)	7.2	(0.5)	23.7	(0.8)	8.0	(0.5)
	Sweden	17.3	(0.9)	15.8	(0.8)	18.3	(1.0)	12.6	(0.7)	21.1	(1.0)	11.4	(0.8)
	Switzerland	14.5	(0.8)	21.2	(1.5)	13.5	(0.9)	22.6	(1.2)	13.5	(0.8)	24.1	(1.4)
	Turkey	52.2	(2.6)	5.5	(1.6)	52.1	(1.8)	4.2	(1.2)	42.1	(1.8)	5.6	(1.2)
	United Kingdom	m	m	m	m	19.8	(0.8)	11.1	(0.6)	20.2	(0.9)	9.8	(0.7)
	United States*	25.7	(1.2)	10.1	(0.7)	28.1	(1.7)	7.6	(0.8)	23.4	(1.3)	9.9	(1.0)
	OECD average-29a	21.7	(0.2)	14.4	(0.2)	21.4	(0.2)	13.1	(0.2)	20.8	(0.2)	13.2	(0.2)
	OECD average-30	21.6	(0.2)	14.4	(0.2)	21.3	(0.2)	13.2	(0.2)	m	m	m	m
	OECD average-36b	m	m	m	m	23.9	(0.2)	12.1	(0.1)	23.5	(0.2)	12.0	(0.1)
	OECD average-37	m	m	m	m	23.8	(0.2)	12.2	(0.1)	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [2/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

	Proficiency level in PISA 2003				Proficiency level in PISA 2006				Proficiency level in PISA 2009			
	Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
Partners	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Albania	m	m	m	m	m	m	m	m	67.7	(1.9)	0.4	(0.2)
Argentina	m	m	m	m	64.1	(2.5)	1.0	(0.4)	63.6	(2.0)	0.9	(0.3)
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	75.2	(1.7)	1.2	(0.4)	72.5	(1.2)	1.0	(0.3)	69.1	(1.2)	0.8	(0.2)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	m	m	m	m	53.3	(2.4)	3.1	(0.8)	47.1	(2.5)	3.8	(1.0)
Costa Rica	m	m	m	m	m	m	m	m	56.7	(1.9)	0.3	(0.2)
Croatia	m	m	m	m	28.6	(1.2)	4.7	(0.5)	33.2	(1.4)	4.9	(0.7)
Cyprus	m	m	m	m	m	m	m	m	m	m	m	m
Dominican Republic	m	m	m	m	m	m	m	m	m	m	m	m
Georgia	m	m	m	m	m	m	m	m	68.7	(1.2)	0.6	(0.2)
Hong Kong (China)*	10.4	(1.2)	30.7	(1.5)	9.5	(0.9)	27.7	(1.2)	8.8	(0.7)	30.7	(1.2)
Indonesia	78.1	(1.7)	0.2	(0.1)	65.8	(3.1)	0.4	(0.2)	76.7	(1.9)	0.1	(0.0)
Jordan	m	m	m	m	66.4	(1.6)	0.2	(0.1)	65.3	(1.9)	0.3	(0.2)
Kazakhstan	m	m	m	m	m	m	m	m	59.1	(1.5)	1.2	(0.4)
Kosovo	m	m	m	m	m	m	m	m	m	m	m	m
Lebanon	m	m	m	m	m	m	m	m	m	m	m	m
Macao (China)	11.2	(1.2)	18.7	(1.4)	10.9	(0.7)	17.4	(0.7)	11.0	(0.5)	17.1	(0.5)
Malaysia	m	m	m	m	m	m	m	m	59.3	(1.6)	0.4	(0.1)
Malta	m	m	m	m	m	m	m	m	33.7	(0.8)	7.7	(0.4)
Moldova	m	m	m	m	m	m	m	m	60.7	(1.6)	0.7	(0.2)
Montenegro	m	m	m	m	60.1	(1.0)	0.8	(0.2)	58.4	(1.1)	1.0	(0.2)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	m	m	m	m	m	m
Panama	m	m	m	m	m	m	m	m	78.8	(2.2)	0.4	(0.2)
Peru	m	m	m	m	m	m	m	m	73.5	(1.8)	0.6	(0.2)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	m	m	m	m	87.2	(0.6)	0.6	(0.1)	73.8	(0.4)	1.8	(0.2)
Romania	m	m	m	m	52.7	(2.2)	1.3	(0.3)	47.0	(2.0)	1.3	(0.3)
Russia	30.2	(1.8)	7.0	(0.8)	26.6	(1.6)	7.4	(0.8)	28.6	(1.5)	5.2	(0.8)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	m	m	m	m	42.6	(1.7)	2.8	(0.4)	40.6	(1.4)	3.5	(0.5)
Singapore	m	m	m	m	m	m	m	m	9.8	(0.6)	35.6	(0.8)
Chinese Taipei	m	m	m	m	12.0	(1.1)	31.9	(1.4)	12.8	(0.8)	28.6	(1.5)
Thailand	54.0	(1.7)	1.6	(0.4)	53.0	(1.3)	1.3	(0.3)	52.5	(1.6)	1.3	(0.4)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	m	m	m	m	m	m	51.3	(1.1)	2.9	(0.3)
Uruguay	48.1	(1.5)	2.8	(0.4)	46.1	(1.2)	3.2	(0.5)	47.6	(1.3)	2.4	(0.4)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [3/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

	Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
	Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD												
Australia	19.7	(0.6)	14.8	(0.6)	22.0	(0.6)	11.3	(0.6)	22.4	(0.7)	10.5	(0.5)
Austria	18.7	(1.0)	14.3	(0.9)	21.8	(1.1)	12.5	(0.9)	21.1	(1.2)	12.6	(0.8)
Belgium	19.0	(0.8)	19.5	(0.8)	20.1	(1.0)	15.9	(0.7)	19.7	(0.9)	15.7	(0.9)
Canada	13.8	(0.5)	16.4	(0.6)	14.4	(0.7)	15.1	(0.8)	16.3	(0.7)	15.3	(0.7)
Chile	51.5	(1.7)	1.6	(0.2)	49.4	(1.3)	1.4	(0.2)	51.9	(1.3)	1.2	(0.2)
Colombia	73.8	(1.4)	0.3	(0.1)	66.3	(1.2)	0.3	(0.1)	65.4	(1.6)	0.5	(0.1)
Czech Republic	21.0	(1.2)	12.9	(0.8)	21.7	(1.1)	10.4	(0.8)	20.4	(1.1)	12.7	(0.7)
Denmark	16.8	(1.0)	10.0	(0.7)	13.6	(0.9)	11.7	(0.7)	14.6	(0.6)	11.6	(0.7)
Estonia	10.5	(0.6)	14.6	(0.8)	11.2	(0.7)	14.2	(0.8)	10.2	(0.6)	15.5	(0.8)
Finland	12.3	(0.7)	15.3	(0.7)	13.6	(0.8)	11.7	(0.7)	15.0	(0.7)	11.1	(0.6)
France	22.4	(0.9)	12.9	(0.8)	23.5	(0.9)	11.4	(0.7)	21.3	(0.8)	11.0	(0.8)
Germany	17.7	(1.0)	17.5	(0.9)	17.2	(1.0)	12.9	(0.8)	21.1	(1.1)	13.3	(0.8)
Greece	35.7	(1.3)	3.9	(0.4)	35.8	(1.8)	3.9	(0.5)	35.8	(1.5)	3.7	(0.5)
Hungary	28.1	(1.3)	9.3	(1.1)	28.0	(1.2)	8.1	(0.6)	25.6	(1.0)	8.0	(0.7)
Iceland	21.5	(0.7)	11.2	(0.7)	23.6	(1.0)	10.3	(0.8)	20.7	(1.0)	10.4	(0.6)
Ireland	16.9	(1.0)	10.7	(0.5)	15.0	(0.9)	9.8	(0.6)	15.7	(0.8)	8.2	(0.7)
Israel	33.5	(1.7)	9.4	(1.0)	32.1	(1.4)	8.9	(0.9)	34.1	(1.4)	8.8	(0.6)
Italy	24.7	(0.8)	9.9	(0.6)	23.3	(1.1)	10.5	(0.8)	23.8	(1.1)	9.5	(0.8)
Japan	11.1	(1.0)	23.7	(1.5)	10.7	(0.8)	20.3	(1.3)	11.5	(0.8)	18.3	(1.1)
Korea	9.1	(0.9)	30.9	(1.8)	15.5	(1.1)	20.9	(1.3)	15.0	(0.9)	21.4	(1.1)
Latvia	19.9	(1.1)	8.0	(0.8)	21.4	(1.0)	5.2	(0.4)	17.3	(1.0)	8.5	(0.6)
Lithuania	26.0	(1.2)	8.1	(0.6)	25.4	(1.1)	6.9	(0.7)	25.6	(0.9)	8.4	(0.5)
Luxembourg	24.3	(0.5)	11.2	(0.4)	25.8	(0.7)	10.0	(0.5)	27.2	(0.7)	10.8	(0.6)
Mexico	54.7	(0.8)	0.6	(0.1)	56.6	(1.3)	0.3	(0.1)	56.2	(1.4)	0.5	(0.1)
Netherlands*	14.8	(1.3)	19.3	(1.2)	16.7	(0.9)	15.5	(0.8)	15.8	(1.1)	18.4	(1.0)
New Zealand	22.6	(0.8)	15.0	(0.9)	21.6	(1.0)	11.4	(0.7)	21.8	(0.8)	11.6	(0.5)
Norway	22.3	(1.1)	9.4	(0.7)	17.1	(0.8)	10.6	(0.7)	18.9	(0.8)	12.2	(0.7)
Poland	14.4	(0.9)	16.7	(1.3)	17.2	(1.0)	12.2	(0.9)	14.7	(0.8)	15.8	(1.0)
Portugal*	24.9	(1.5)	10.6	(0.8)	23.8	(1.0)	11.4	(0.7)	23.3	(1.0)	11.6	(0.7)
Slovak Republic	27.5	(1.3)	11.0	(0.9)	27.7	(1.2)	7.8	(0.6)	25.1	(1.1)	10.7	(0.7)
Slovenia	20.1	(0.6)	13.7	(0.6)	16.1	(0.6)	13.5	(0.7)	16.4	(0.6)	13.6	(0.7)
Spain	23.6	(0.8)	8.0	(0.4)	22.2	(1.0)	7.2	(0.6)	24.7	(0.6)	7.3	(0.4)
Sweden	27.1	(1.1)	8.0	(0.5)	20.8	(1.2)	10.4	(0.9)	18.8	(1.0)	12.6	(0.8)
Switzerland	12.4	(0.7)	21.4	(1.2)	15.8	(1.0)	19.2	(1.0)	16.8	(0.9)	17.0	(1.0)
Turkey	42.0	(1.9)	5.9	(1.1)	51.4	(2.2)	1.1	(0.4)	36.7	(1.1)	4.8	(0.6)
United Kingdom	21.8	(1.3)	11.8	(0.8)	21.9	(1.0)	10.6	(0.7)	19.2	(0.9)	12.9	(0.8)
United States*	25.8	(1.4)	8.8	(0.8)	29.4	(1.4)	5.9	(0.7)	27.1	(1.4)	8.3	(0.8)
OECD average-29a	22.3	(0.2)	12.8	(0.2)	22.9	(0.2)	10.8	(0.1)	22.2	(0.2)	11.4	(0.1)
OECD average-30	22.2	(0.2)	12.9	(0.2)	22.9	(0.2)	10.8	(0.1)	22.1	(0.2)	11.4	(0.1)
OECD average-36b	24.5	(0.2)	12.0	(0.1)	24.7	(0.2)	10.2	(0.1)	24.1	(0.2)	10.9	(0.1)
OECD average-37	24.4	(0.2)	12.1	(0.1)	24.6	(0.2)	10.3	(0.1)	24.0	(0.2)	10.9	(0.1)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [4/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

	Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
	Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
Partners	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Albania	60.7	(1.0)	0.8	(0.2)	53.3	(1.9)	1.1	(0.2)	42.4	(1.4)	2.3	(0.3)
Argentina	66.5	(2.0)	0.3	(0.1)	m	m	m	m	69.0	(1.3)	0.3	(0.1)
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	50.7	(1.3)	2.0	(0.3)
Belarus	m	m	m	m	m	m	m	m	29.4	(1.1)	7.3	(0.6)
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	57.6	(1.6)	0.8	(0.2)
Brazil	68.3	(1.0)	0.7	(0.2)	70.3	(1.2)	0.9	(0.2)	68.1	(1.0)	0.9	(0.2)
Brunei Darussalam	m	m	m	m	m	m	m	m	47.9	(0.7)	3.0	(0.3)
B-S-J-Z (China)	m	m	m	m	m	m	m	m	2.4	(0.4)	44.3	(1.3)
Bulgaria	43.8	(1.8)	4.1	(0.6)	42.1	(1.8)	4.4	(0.6)	44.4	(1.7)	4.2	(0.6)
Costa Rica	59.9	(1.9)	0.6	(0.2)	62.5	(1.5)	0.3	(0.1)	60.0	(1.9)	0.3	(0.1)
Croatia	29.9	(1.4)	7.0	(1.1)	32.0	(1.4)	5.6	(0.5)	31.2	(1.3)	5.1	(0.5)
Cyprus	42.0	(0.6)	3.7	(0.3)	42.6	(0.8)	3.2	(0.4)	36.9	(0.7)	4.4	(0.4)
Dominican Republic	m	m	m	m	90.5	(1.0)	0.0	(0.0)	90.6	(1.0)	0.0	(0.0)
Georgia	m	m	m	m	57.1	(1.2)	1.6	(0.4)	61.1	(1.3)	1.0	(0.3)
Hong Kong (China)*	8.5	(0.8)	33.7	(1.4)	9.0	(0.8)	26.5	(1.1)	9.2	(0.8)	29.0	(1.1)
Indonesia	75.7	(2.1)	0.3	(0.2)	68.6	(1.6)	0.7	(0.2)	71.9	(1.5)	0.5	(0.2)
Jordan	68.6	(1.5)	0.6	(0.4)	67.5	(1.3)	0.3	(0.1)	59.3	(1.6)	0.7	(0.2)
Kazakhstan	45.2	(1.7)	0.9	(0.3)	m	m	m	m	49.1	(0.9)	1.9	(0.2)
Kosovo	m	m	m	m	77.7	(1.0)	0.0	(0.0)	76.6	(0.9)	0.1	(0.1)
Lebanon	m	m	m	m	60.2	(1.6)	2.0	(0.3)	59.8	(1.7)	2.0	(0.3)
Macao (China)	10.8	(0.5)	24.3	(0.6)	6.6	(0.5)	21.9	(0.6)	5.0	(0.5)	27.6	(0.8)
Malaysia	51.8	(1.7)	1.3	(0.3)	m	m	m	m	41.5	(1.4)	2.5	(0.4)
Malta	m	m	m	m	29.1	(0.8)	11.8	(0.7)	30.2	(1.0)	8.5	(0.7)
Moldova	m	m	m	m	50.3	(1.2)	1.7	(0.3)	50.3	(1.1)	2.4	(0.4)
Montenegro	56.6	(1.0)	1.0	(0.2)	51.9	(1.0)	1.5	(0.2)	46.2	(0.8)	1.8	(0.2)
Morocco	m	m	m	m	m	m	m	m	75.6	(1.6)	0.1	(0.1)
North Macedonia	m	m	m	m	70.2	(0.8)	0.8	(0.2)	61.0	(0.9)	1.1	(0.2)
Panama	m	m	m	m	m	m	m	m	81.2	(1.3)	0.1	(0.1)
Peru	74.6	(1.8)	0.6	(0.2)	66.2	(1.4)	0.4	(0.1)	60.3	(1.3)	0.9	(0.2)
Philippines	m	m	m	m	m	m	m	m	80.7	(1.6)	0.1	(0.1)
Qatar	69.6	(0.5)	2.0	(0.2)	58.7	(0.7)	2.2	(0.2)	53.7	(0.6)	2.9	(0.2)
Romania	40.8	(1.9)	3.2	(0.6)	39.9	(1.8)	3.3	(0.5)	46.6	(2.3)	3.2	(0.6)
Russia	24.0	(1.1)	7.8	(0.8)	18.9	(1.2)	8.8	(0.7)	21.6	(1.3)	8.1	(0.7)
Saudi Arabia	m	m	m	m	m	m	m	m	72.7	(1.5)	0.2	(0.1)
Serbia	38.9	(1.5)	4.6	(0.7)	m	m	m	m	39.7	(1.4)	5.2	(0.4)
Singapore	8.3	(0.5)	40.0	(0.7)	7.6	(0.4)	34.8	(0.8)	7.1	(0.4)	36.9	(0.8)
Chinese Taipei	12.8	(0.8)	37.2	(1.2)	12.7	(0.7)	28.1	(1.2)	14.0	(0.8)	23.2	(1.1)
Thailand	49.7	(1.7)	2.6	(0.5)	53.8	(1.6)	1.4	(0.3)	52.7	(1.7)	2.3	(0.4)
Ukraine	m	m	m	m	m	m	m	m	35.9	(1.6)	5.0	(0.6)
United Arab Emirates	46.3	(1.2)	3.5	(0.3)	48.7	(1.2)	3.7	(0.3)	45.5	(0.9)	5.4	(0.3)
Uruguay	55.8	(1.3)	1.4	(0.3)	52.4	(1.2)	1.7	(0.4)	50.7	(1.5)	1.0	(0.3)
Viet Nam**	14.2	(1.7)	13.3	(1.5)	19.1	(1.7)	9.3	(1.3)	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [5/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

		Change between 2003 and 2018 (PISA 2018 - PISA 2003)				Change between 2006 and 2018 (PISA 2018 - PISA 2006)				Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
		% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
OECD	Australia	**8.1**	(1.2)	**-9.3**	(1.0)	**9.5**	(1.3)	**-6.0**	(1.1)	**6.6**	(1.6)	**-6.0**	(1.2)	2.8	(1.4)	**-4.3**	(1.0)	0.5	(1.0)	-0.9	(0.9)
	Austria	2.3	(1.8)	-1.7	(1.4)	1.1	(2.0)	**-3.2**	(1.4)	m	m	m	m	2.4	(1.8)	-1.7	(1.5)	-0.7	(1.7)	0.1	(1.2)
	Belgium	**3.2**	(1.3)	**-10.7**	(1.4)	2.3	(1.5)	**-6.6**	(1.5)	0.6	(1.5)	**-4.6**	(1.6)	0.7	(1.4)	**-3.8**	(1.5)	-0.4	(1.4)	-0.1	(1.2)
	Canada	**6.1**	(1.1)	**-5.0**	(1.3)	**5.5**	(1.3)	-2.6	(1.4)	**4.8**	(1.4)	**-3.0**	(1.5)	2.4	(1.3)	-1.1	(1.4)	1.9	(1.1)	0.3	(1.2)
	Chile	m	m	m	m	-3.2	(3.7)	-0.2	(0.4)	0.9	(3.8)	-0.1	(0.4)	0.4	(3.5)	-0.4	(0.3)	2.6	(2.3)	-0.2	(0.3)
	Colombia	m	m	m	m	-6.5	(3.4)	0.1	(0.2)	-5.1	(3.8)	**0.4**	(0.2)	**-8.4**	(3.5)	0.2	(0.2)	-0.9	(2.4)	0.2	(0.2)
	Czech Republic	**3.8**	(1.9)	**-5.6**	(1.4)	1.2	(1.9)	**-5.6**	(1.5)	-1.9	(2.0)	1.0	(1.4)	-0.6	(2.0)	-0.2	(1.3)	-1.3	(1.6)	**2.3**	(1.1)
	Denmark	-0.8	(1.2)	**-4.3**	(1.3)	0.9	(1.3)	-2.1	(1.2)	-2.5	(1.4)	0.1	(1.4)	-2.3	(1.4)	1.7	(1.2)	1.0	(1.1)	0.0	(1.1)
	Estonia	m	m	m	m	-1.9	(1.4)	2.9	(1.5)	-2.4	(1.4)	**3.4**	(1.7)	-0.3	(1.2)	0.9	(1.6)	-1.0	(1.1)	1.3	(1.2)
	Finland	**8.2**	(1.1)	**-12.3**	(1.1)	**9.0**	(1.3)	**-13.3**	(1.3)	**7.1**	(1.4)	**-10.5**	(1.3)	**2.7**	(1.4)	**-4.1**	(1.2)	1.4	(1.2)	-0.6	(1.0)
	France	**4.6**	(1.5)	**-4.1**	(1.3)	-1.0	(1.7)	-1.5	(1.3)	-1.3	(1.7)	-2.7	(1.4)	-1.1	(1.4)	-1.9	(1.2)	-2.2	(1.3)	-0.4	(1.1)
	Germany	-0.5	(1.7)	**-2.9**	(1.4)	1.2	(1.9)	-2.1	(1.5)	2.5	(1.8)	**-4.5**	(1.6)	**3.4**	(1.7)	**-4.1**	(1.5)	**3.9**	(1.5)	0.4	(1.2)
	Greece	-3.1	(2.7)	-0.3	(0.8)	3.5	(2.6)	-1.3	(0.7)	5.5	(3.1)	**-2.0**	(0.8)	0.1	(2.7)	-0.2	(0.7)	0.1	(2.5)	-0.2	(0.7)
	Hungary	2.6	(1.8)	**-2.7**	(1.2)	**4.5**	(2.0)	-2.4	(1.2)	3.3	(2.5)	-2.1	(1.4)	-2.4	(2.2)	-1.3	(1.4)	-2.3	(1.7)	-0.2	(1.0)
	Iceland	**5.7**	(1.4)	**-5.1**	(1.0)	**3.9**	(1.6)	**-2.3**	(1.1)	**3.7**	(1.7)	**-3.2**	(1.1)	-0.8	(1.6)	-0.8	(1.1)	**-2.9**	(1.5)	0.1	(1.1)
	Ireland	-1.1	(1.5)	**-3.1**	(1.1)	-0.7	(1.8)	-2.0	(1.1)	**-5.1**	(1.8)	1.6	(1.1)	-1.2	(1.7)	**-2.4**	(1.0)	0.7	(1.4)	-1.6	(0.9)
	Israel	m	m	m	m	**-7.9**	(2.4)	**2.7**	(0.9)	**-5.4**	(2.2)	**2.9**	(1.0)	0.6	(2.4)	-0.6	(1.3)	2.0	(2.0)	-0.1	(1.1)
	Italy	**-8.1**	(2.1)	**2.5**	(1.0)	**-9.0**	(1.9)	**3.3**	(1.1)	-1.1	(1.9)	0.6	(1.1)	-0.8	(1.9)	-0.4	(1.1)	0.6	(1.7)	-1.0	(1.2)
	Japan	-1.9	(1.4)	**-5.9**	(2.1)	-1.6	(1.4)	0.0	(1.9)	-1.0	(1.4)	-2.5	(2.2)	0.4	(1.3)	**-5.3**	(2.2)	0.8	(1.1)	-2.0	(1.8)
	Korea	**5.5**	(1.3)	-3.4	(2.0)	**6.1**	(1.4)	**-5.7**	(2.3)	**6.9**	(1.5)	-4.2	(2.5)	**5.9**	(1.4)	**-9.5**	(2.5)	-0.5	(1.4)	0.5	(1.9)
	Latvia	**-6.4**	(1.9)	0.5	(1.1)	-3.4	(1.9)	**1.9**	(1.0)	**-5.2**	(2.1)	**2.8**	(1.0)	-2.6	(1.9)	0.5	(1.1)	**-4.1**	(1.5)	**3.3**	(0.8)
	Lithuania	m	m	m	m	2.7	(1.9)	-0.7	(1.1)	-0.7	(2.1)	1.5	(0.9)	-0.4	(2.0)	0.4	(0.9)	0.2	(1.6)	1.5	(0.9)
	Luxembourg	**5.5**	(1.3)	0.0	(0.9)	**4.4**	(1.4)	0.2	(0.9)	**3.3**	(1.6)	-0.5	(1.0)	2.9	(1.5)	-0.4	(0.8)	1.4	(1.1)	0.8	(0.8)
	Mexico	**-9.7**	(2.9)	0.1	(0.2)	-0.3	(3.2)	-0.3	(0.2)	5.4	(3.6)	-0.2	(0.2)	1.5	(3.2)	-0.1	(0.1)	-0.4	(2.3)	0.2	(0.2)
	Netherlands*	**4.8**	(1.6)	**-7.1**	(1.8)	**4.2**	(1.6)	-2.7	(1.8)	2.3	(1.9)	-1.4	(2.2)	1.0	(1.7)	-0.8	(1.9)	-1.0	(1.4)	**2.9**	(1.4)
	New Zealand	**6.7**	(1.5)	**-9.1**	(1.0)	**7.7**	(1.6)	**-7.4**	(1.2)	**6.4**	(1.8)	**-7.3**	(1.3)	-0.9	(1.7)	**-3.4**	(1.2)	0.1	(1.4)	0.2	(1.0)
	Norway	-2.0	(1.4)	0.8	(1.1)	**-3.4**	(1.6)	1.8	(1.2)	0.7	(1.5)	2.0	(1.3)	**-3.4**	(1.6)	**2.8**	(1.2)	1.8	(1.2)	1.5	(1.0)
	Poland	**-7.3**	(1.4)	**5.7**	(1.3)	**-5.1**	(1.3)	**5.2**	(1.5)	**-5.8**	(1.5)	**5.4**	(1.6)	0.3	(1.3)	-1.0	(1.9)	-2.5	(1.3)	**3.6**	(1.4)
	Portugal*	**-6.8**	(2.1)	**6.2**	(1.0)	**-7.4**	(2.0)	**5.9**	(1.0)	-0.4	(1.8)	2.0	(1.3)	-1.6	(2.0)	1.0	(1.2)	-0.5	(1.5)	0.2	(1.1)
	Slovak Republic	**5.1**	(1.9)	-1.9	(1.2)	**4.2**	(1.7)	-0.2	(1.3)	**4.1**	(1.9)	-1.9	(1.4)	-2.4	(1.9)	-0.2	(1.3)	-2.6	(1.7)	**2.9**	(1.0)
	Slovenia	m	m	m	m	-1.2	(1.1)	-0.1	(1.2)	**-3.9**	(1.1)	-0.6	(1.3)	**-3.7**	(1.1)	-0.1	(1.2)	0.3	(0.9)	0.1	(1.1)
	Spain	1.7	(1.5)	-0.7	(0.8)	0.0	(1.8)	0.0	(0.8)	1.0	(1.9)	-0.8	(0.8)	1.1	(1.7)	-0.7	(0.7)	2.5	(1.3)	0.0	(0.8)
	Sweden	1.5	(1.5)	**-3.2**	(1.2)	0.5	(1.6)	0.0	(1.2)	-2.3	(1.7)	1.2	(1.4)	**-8.3**	(1.7)	**4.6**	(1.1)	-2.0	(1.6)	2.2	(1.2)
	Switzerland	2.3	(1.3)	**-4.3**	(1.9)	**3.3**	(1.4)	**-5.7**	(1.7)	**3.4**	(1.5)	**-7.2**	(1.9)	**4.4**	(1.4)	**-4.4**	(1.7)	1.0	(1.4)	-2.3	(1.5)
	Turkey	**-15.6**	(3.3)	-0.7	(1.7)	**-15.4**	(3.1)	0.6	(1.3)	-5.5	(3.6)	-0.9	(1.4)	-5.3	(3.4)	-1.1	(1.3)	**-14.7**	(2.8)	**3.6**	(0.7)
	United Kingdom	m	m	m	m	-0.5	(1.6)	1.7	(1.1)	-1.0	(1.8)	**3.0**	(1.2)	-2.6	(1.9)	1.0	(1.3)	-2.6	(1.5)	**2.2**	(1.1)
	United States*	1.4	(2.1)	-1.8	(1.1)	-1.0	(2.6)	0.6	(1.2)	3.7	(2.5)	-1.7	(1.3)	1.3	(2.4)	-0.5	(1.2)	-2.3	(2.1)	**2.4**	(1.1)
	OECD average-29a	0.5	(0.8)	**-3.0**	(0.5)	0.8	(1.0)	**-1.7**	(0.6)	1.3	(1.2)	**-1.7**	(0.7)	-0.1	(1.1)	**-1.4**	(0.6)	-0.8	(0.6)	0.6	(0.4)
	OECD average-30	0.5	(0.8)	**-3.0**	(0.5)	0.8	(1.0)	**-1.8**	(0.6)	m	m	m	m	0.0	(1.1)	**-1.4**	(0.6)	-0.8	(0.6)	0.6	(0.4)
	OECD average-36b	m	m	m	m	0.1	(1.0)	**-1.2**	(0.5)	0.6	(1.3)	-1.1	(0.6)	-0.5	(1.1)	-1.1	(0.6)	-0.6	(0.6)	**0.6**	(0.3)
	OECD average-37	m	m	m	m	0.2	(1.0)	**-1.3**	(0.5)	m	m	m	m	-0.4	(1.1)	**-1.1**	(0.6)	-0.6	(0.6)	**0.6**	(0.3)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.8 [6/6] **Percentage of low achievers and top performers in mathematics, 2003 through 2018**

	Change between 2003 and 2018 (PISA 2018 - PISA 2003)				Change between 2006 and 2018 (PISA 2018 - PISA 2006)				Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
	Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)		Below Level 2 (less than 420.07 score points)		Level 5 or above (at or above 606.99 score points)	
	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
Partners																				
Albania	m	m	m	m	m	m	m	m	**-25.4**	(3.8)	**1.9**	(0.3)	**-18.3**	(3.1)	**1.5**	(0.3)	**-10.9**	(2.7)	**1.2**	(0.4)
Argentina	m	m	m	m	4.8	(3.3)	-0.7	(0.4)	5.4	(3.1)	-0.5	(0.3)	2.5	(3.0)	0.1	(0.1)	m	m	m	m
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	**-7.1**	(2.2)	-0.3	(0.5)	**-4.4**	(2.1)	-0.1	(0.4)	-1.0	(2.3)	0.1	(0.3)	-0.2	(2.0)	0.2	(0.3)	-2.2	(1.7)	0.0	(0.3)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	m	m	m	m	**-8.9**	(3.3)	1.1	(1.0)	-2.7	(3.6)	0.4	(1.2)	0.7	(3.0)	0.2	(0.9)	2.3	(2.6)	-0.2	(0.9)
Costa Rica	m	m	m	m	m	m	m	m	3.3	(4.3)	0.0	(0.2)	0.1	(4.0)	-0.2	(0.2)	-2.5	(2.8)	0.1	(0.2)
Croatia	m	m	m	m	2.6	(2.5)	0.4	(0.8)	-2.0	(2.9)	0.2	(0.9)	1.3	(2.7)	-1.8	(1.3)	-0.9	(2.1)	-0.4	(0.8)
Cyprus	m	m	m	m	m	m	m	m	m	m	m	m	**-5.2**	(1.9)	0.7	(0.5)	**-5.7**	(1.4)	**1.2**	(0.6)
Dominican Republic	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	0.0	(1.4)	0.0	(0.0)
Georgia	m	m	m	m	m	m	m	m	**-7.7**	(2.9)	0.4	(0.4)	m	m	m	m	**4.0**	(2.0)	-0.5	(0.5)
Hong Kong (China)*	-1.2	(1.4)	-1.7	(2.1)	-0.4	(1.3)	1.2	(2.1)	0.4	(1.1)	-1.7	(2.3)	0.7	(1.2)	**-4.7**	(2.3)	0.2	(1.1)	2.4	(1.7)
Indonesia	**-6.3**	(2.8)	0.2	(0.2)	6.1	(4.1)	0.1	(0.2)	-4.8	(3.6)	**0.4**	(0.2)	-3.8	(3.5)	0.2	(0.2)	3.2	(2.5)	-0.2	(0.2)
Jordan	m	m	m	m	**-7.0**	(3.4)	0.4	(0.3)	-5.9	(4.0)	0.4	(0.3)	**-9.2**	(3.5)	0.1	(0.5)	**-8.2**	(2.5)	0.4	(0.2)
Kazakhstan	m	m	m	m	m	m	m	m	**-10.0**	(3.4)	0.8	(0.5)	3.9	(3.3)	**1.0**	(0.4)	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	-1.1	(1.7)	0.1	(0.1)
Lebanon	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	-0.5	(2.4)	0.1	(0.4)
Macao (China)	**-6.2**	(1.3)	**9.0**	(2.5)	**-5.9**	(0.8)	**10.2**	(2.7)	**-6.0**	(0.7)	**10.5**	(3.1)	**-5.8**	(0.7)	3.3	(2.9)	**-1.6**	(0.7)	**5.8**	(1.7)
Malaysia	m	m	m	m	m	m	m	m	**-17.8**	(3.8)	**2.1**	(0.5)	**-10.3**	(3.6)	**1.1**	(0.5)	m	m	m	m
Malta	m	m	m	m	m	m	m	m	-3.5	(1.8)	0.8	(0.9)	m	m	m	m	1.1	(1.4)	**-3.4**	(1.0)
Moldova	m	m	m	m	m	m	m	m	**-10.3**	(2.7)	**1.8**	(0.4)	m	m	m	m	0.1	(1.9)	0.7	(0.5)
Montenegro	m	m	m	m	**-13.9**	(2.5)	**0.9**	(0.3)	**-12.3**	(3.0)	**0.8**	(0.3)	**-10.5**	(2.7)	**0.7**	(0.3)	**-5.7**	(1.7)	0.3	(0.3)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	**-9.2**	(1.6)	0.3	(0.3)
Panama	m	m	m	m	m	m	m	m	2.4	(2.9)	-0.3	(0.2)	m	m	m	m	m	m	m	m
Peru	m	m	m	m	m	m	m	m	**-13.2**	(3.5)	0.2	(0.3)	**-14.2**	(3.2)	0.3	(0.3)	**-5.8**	(2.2)	**0.5**	(0.2)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	m	m	m	m	**-33.5**	(1.8)	**2.3**	(0.2)	**-20.1**	(2.1)	**1.2**	(0.3)	**-15.9**	(1.9)	**0.9**	(0.3)	**-5.0**	(1.3)	**0.7**	(0.3)
Romania	m	m	m	m	-6.2	(3.5)	**1.9**	(0.7)	-0.5	(3.6)	**1.9**	(0.7)	5.7	(3.4)	0.0	(0.9)	**6.6**	(3.0)	-0.1	(0.8)
Russia	**-8.6**	(2.4)	1.0	(1.1)	**-5.0**	(2.4)	0.6	(1.1)	**-6.9**	(2.5)	**2.8**	(1.1)	-2.3	(2.2)	0.3	(1.2)	2.7	(1.8)	-0.7	(1.0)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	m	m	m	m	-2.9	(2.5)	**2.4**	(0.6)	-0.8	(2.6)	**1.7**	(0.7)	0.8	(2.5)	0.7	(0.9)	m	m	m	m
Singapore	m	m	m	m	m	m	m	m	**-2.7**	(0.7)	1.3	(2.7)	-1.1	(0.7)	-3.1	(2.4)	-0.4	(0.6)	2.1	(1.6)
Chinese Taipei	m	m	m	m	2.0	(1.4)	**-8.7**	(2.2)	1.2	(1.1)	**-5.4**	(2.4)	1.1	(1.2)	**-14.0**	(2.1)	1.3	(1.1)	**-4.9**	(1.8)
Thailand	-1.3	(2.8)	0.6	(0.5)	-0.3	(2.9)	**0.9**	(0.4)	0.2	(3.4)	1.0	(0.6)	3.0	(3.3)	-0.3	(0.6)	-1.1	(2.6)	0.8	(0.5)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	m	m	m	m	m	m	**-5.8**	(2.3)	**2.6**	(0.5)	-0.8	(2.2)	**2.0**	(0.5)	-3.2	(1.7)	**1.7**	(0.5)
Uruguay	2.6	(2.7)	**-1.8**	(0.5)	4.6	(2.8)	**-2.2**	(0.5)	3.1	(3.2)	**-1.4**	(0.4)	-5.1	(3.0)	-0.3	(0.4)	-1.7	(2.3)	-0.7	(0.5)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.9 [1/4] **Percentage of low achievers and top performers in science, 2006 through 2018**

		Proficiency level in PISA 2006				Proficiency level in PISA 2009				Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	**Australia**	12.9	(0.6)	14.6	(0.7)	12.6	(0.6)	14.5	(0.8)	13.6	(0.5)	13.6	(0.5)	17.6	(0.6)	11.2	(0.5)	18.9	(0.6)	9.5	(0.5)
	Austria	16.3	(1.4)	10.0	(0.8)	m	m	m	m	15.8	(1.0)	7.9	(0.7)	20.8	(1.0)	7.7	(0.5)	21.9	(1.0)	6.3	(0.6)
	Belgium	17.0	(1.0)	10.1	(0.5)	18.0	(0.8)	10.1	(0.7)	17.7	(0.9)	9.1	(0.4)	19.8	(0.9)	9.0	(0.4)	20.0	(0.9)	8.0	(0.5)
	Canada	10.0	(0.6)	14.4	(0.5)	9.6	(0.5)	12.1	(0.5)	10.4	(0.5)	11.3	(0.5)	11.1	(0.5)	12.4	(0.6)	13.4	(0.5)	11.3	(0.6)
	Chile	39.7	(2.1)	1.9	(0.3)	32.3	(1.4)	1.1	(0.2)	34.5	(1.6)	1.0	(0.2)	34.8	(1.2)	1.2	(0.2)	35.3	(1.2)	1.0	(0.2)
	Colombia	60.2	(1.8)	0.2	(0.1)	54.1	(1.9)	0.1	(0.1)	56.2	(1.6)	0.1	(0.1)	49.0	(1.3)	0.4	(0.1)	50.4	(1.7)	0.4	(0.1)
	Czech Republic	15.5	(1.2)	11.6	(0.9)	17.3	(1.2)	8.4	(0.7)	13.8	(1.1)	7.6	(0.6)	20.7	(1.0)	7.3	(0.5)	18.8	(1.1)	7.5	(0.5)
	Denmark	18.4	(1.1)	6.8	(0.7)	16.6	(0.8)	6.7	(0.6)	16.7	(1.0)	6.8	(0.7)	15.9	(0.8)	7.0	(0.6)	18.7	(0.7)	5.5	(0.5)
	Estonia	7.7	(0.6)	11.5	(0.8)	8.3	(0.8)	10.4	(0.8)	5.0	(0.5)	12.8	(0.7)	8.8	(0.7)	13.5	(0.7)	8.8	(0.6)	12.2	(0.6)
	Finland	4.1	(0.5)	20.9	(0.8)	6.0	(0.5)	18.7	(0.9)	7.7	(0.6)	17.1	(0.7)	11.5	(0.7)	14.3	(0.6)	12.9	(0.7)	12.3	(0.7)
	France	21.2	(1.4)	8.0	(0.7)	19.3	(1.3)	8.1	(0.8)	18.7	(1.0)	7.9	(0.8)	22.1	(0.9)	8.0	(0.5)	20.5	(0.8)	6.6	(0.5)
	Germany	15.4	(1.3)	11.8	(0.7)	14.8	(1.0)	12.8	(0.8)	12.2	(0.9)	12.2	(1.0)	17.0	(1.0)	10.6	(0.6)	19.6	(1.0)	10.0	(0.6)
	Greece	24.0	(1.3)	3.4	(0.4)	25.3	(1.6)	3.1	(0.4)	25.5	(1.5)	2.5	(0.4)	32.7	(1.9)	2.1	(0.3)	31.7	(1.5)	1.3	(0.2)
	Hungary	15.0	(1.0)	6.9	(0.6)	14.1	(1.4)	5.4	(0.6)	18.0	(1.1)	5.9	(0.8)	26.0	(1.0)	4.6	(0.5)	24.1	(0.9)	4.7	(0.5)
	Iceland	20.6	(0.8)	6.3	(0.5)	17.9	(0.7)	7.0	(0.4)	24.0	(0.8)	5.2	(0.6)	25.3	(0.9)	3.8	(0.4)	25.0	(0.9)	3.8	(0.4)
	Ireland	15.5	(1.1)	9.4	(0.7)	15.2	(1.1)	8.7	(0.8)	11.1	(0.9)	10.7	(0.6)	15.3	(1.0)	7.1	(0.5)	17.0	(0.8)	5.8	(0.6)
	Israel	36.1	(1.4)	5.2	(0.6)	33.1	(1.2)	3.9	(0.4)	28.9	(1.7)	5.8	(0.6)	31.4	(1.4)	5.8	(0.5)	33.1	(1.4)	5.8	(0.5)
	Italy	25.3	(0.9)	4.6	(0.3)	20.6	(0.6)	5.8	(0.3)	18.7	(0.7)	6.1	(0.4)	23.2	(1.0)	4.1	(0.4)	25.9	(1.0)	2.7	(0.4)
	Japan	12.0	(1.0)	15.1	(0.8)	10.7	(1.0)	16.9	(0.9)	8.5	(0.9)	18.2	(1.2)	9.6	(0.7)	15.3	(1.0)	10.8	(0.8)	13.1	(0.9)
	Korea	11.2	(1.1)	10.3	(1.1)	6.3	(0.8)	11.6	(1.1)	6.6	(0.8)	11.7	(1.1)	14.4	(0.9)	10.6	(0.8)	14.2	(0.8)	11.8	(0.8)
	Latvia	17.4	(1.2)	4.1	(0.4)	14.7	(1.2)	3.1	(0.5)	12.4	(1.0)	4.4	(0.5)	17.2	(0.8)	3.8	(0.4)	18.5	(0.8)	3.7	(0.4)
	Lithuania	20.3	(1.0)	5.0	(0.7)	17.0	(1.1)	4.6	(0.5)	16.1	(1.1)	5.1	(0.5)	24.7	(1.1)	4.2	(0.5)	22.2	(0.9)	4.4	(0.3)
	Luxembourg	22.1	(0.5)	5.9	(0.4)	23.7	(0.8)	6.7	(0.5)	22.2	(0.6)	8.2	(0.5)	25.9	(0.7)	6.9	(0.4)	26.8	(0.6)	5.4	(0.5)
	Mexico	50.9	(1.4)	0.3	(0.1)	47.4	(1.0)	0.2	(0.0)	47.0	(0.8)	0.1	(0.0)	47.8	(1.3)	0.1	(0.1)	46.8	(1.4)	0.3	(0.1)
	Netherlands*	13.0	(1.0)	13.1	(0.9)	13.2	(1.6)	12.7	(1.2)	13.1	(1.1)	11.8	(1.1)	18.5	(1.0)	11.1	(0.6)	20.0	(1.1)	10.6	(0.8)
	New Zealand	13.7	(0.7)	17.6	(0.8)	13.4	(0.7)	17.6	(0.8)	16.3	(0.9)	13.4	(0.7)	17.4	(0.9)	12.8	(0.7)	18.0	(0.8)	11.3	(0.6)
	Norway	21.1	(1.3)	6.1	(0.5)	15.8	(0.9)	6.4	(0.6)	19.6	(1.1)	7.5	(0.6)	18.7	(0.8)	8.0	(0.5)	20.8	(1.0)	6.8	(0.5)
	Poland	17.0	(0.8)	6.8	(0.5)	13.1	(0.8)	7.5	(0.5)	9.0	(0.7)	10.8	(1.0)	16.3	(0.8)	7.3	(0.6)	13.8	(0.8)	9.3	(0.8)
	Portugal*	24.5	(1.4)	3.1	(0.4)	16.5	(1.1)	4.2	(0.5)	19.0	(1.4)	4.5	(0.5)	17.4	(0.9)	7.4	(0.5)	19.6	(1.0)	5.6	(0.6)
	Slovak Republic	20.2	(1.0)	5.8	(0.5)	19.3	(1.2)	6.2	(0.6)	26.9	(1.6)	4.9	(0.7)	30.7	(1.1)	3.6	(0.4)	29.3	(1.0)	3.7	(0.4)
	Slovenia	13.9	(0.6)	12.9	(0.6)	14.8	(0.5)	9.9	(0.6)	12.9	(0.6)	9.6	(0.7)	15.0	(0.5)	10.6	(0.6)	14.6	(0.7)	7.3	(0.6)
	Spain	19.6	(0.9)	4.9	(0.4)	18.2	(0.9)	4.0	(0.3)	15.7	(0.7)	4.8	(0.3)	18.3	(0.8)	5.0	(0.4)	21.3	(0.6)	4.2	(0.3)
	Sweden	16.4	(0.8)	7.9	(0.5)	19.1	(1.0)	8.1	(0.6)	22.2	(1.1)	6.3	(0.5)	21.6	(1.1)	8.5	(0.7)	19.0	(1.1)	8.3	(0.6)
	Switzerland	16.1	(0.9)	10.5	(0.8)	14.0	(0.8)	10.7	(0.9)	12.8	(0.7)	9.3	(0.8)	18.5	(1.1)	9.8	(0.6)	20.2	(1.0)	7.8	(0.7)
	Turkey	46.6	(1.6)	0.9	(0.3)	30.0	(1.5)	1.1	(0.3)	26.4	(1.5)	1.8	(0.4)	44.5	(2.1)	0.3	(0.1)	25.2	(1.1)	2.5	(0.5)
	United Kingdom	16.7	(0.8)	13.7	(0.6)	15.0	(0.8)	11.4	(0.7)	15.0	(1.1)	11.2	(0.8)	17.4	(0.8)	10.9	(0.7)	17.4	(0.9)	9.7	(0.6)
	United States*	24.4	(1.6)	9.1	(0.7)	18.1	(1.1)	9.2	(1.0)	18.1	(1.3)	7.5	(0.7)	20.3	(1.1)	8.5	(0.6)	18.6	(1.2)	9.1	(0.7)
	OECD average-36b	21.0	(0.2)	8.4	(0.1)	18.8	(0.2)	8.0	(0.1)	18.7	(0.2)	8.0	(0.1)	22.1	(0.2)	7.4	(0.1)	22.0	(0.2)	6.8	(0.1)
	OECD average-37	20.9	(0.2)	8.4	(0.1)	m	m	m	m	18.6	(0.2)	8.0	(0.1)	22.1	(0.2)	7.4	(0.1)	22.0	(0.2)	6.8	(0.1)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.9 [2/4] **Percentage of low achievers and top performers in science, 2006 through 2018**

	Proficiency level in PISA 2006				Proficiency level in PISA 2009				Proficiency level in PISA 2012				Proficiency level in PISA 2015				Proficiency level in PISA 2018			
	Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
Partners																				
Albania	m	m	m	m	57.3	(2.0)	0.1	(0.1)	53.1	(1.2)	0.4	(0.1)	41.7	(1.7)	0.4	(0.2)	47.0	(1.3)	0.2	(0.1)
Argentina	56.3	(2.5)	0.4	(0.1)	52.4	(1.9)	0.7	(0.2)	50.9	(2.2)	0.2	(0.1)	m	m	m	m	53.5	(1.4)	0.5	(0.1)
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	57.8	(1.2)	0.1	(0.1)
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	24.2	(1.2)	2.6	(0.4)
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	56.8	(1.6)	0.1	(0.1)
Brazil	61.0	(1.4)	0.6	(0.2)	54.2	(1.3)	0.6	(0.1)	55.2	(1.1)	0.3	(0.1)	56.6	(1.1)	0.7	(0.1)	55.4	(1.0)	0.8	(0.2)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	45.7	(0.6)	2.3	(0.3)
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	2.1	(0.3)	31.5	(1.3)
Bulgaria	42.6	(2.4)	3.1	(0.6)	38.8	(2.5)	2.6	(0.5)	36.9	(2.0)	3.1	(0.6)	37.9	(1.9)	2.9	(0.4)	46.5	(1.6)	1.5	(0.3)
Costa Rica	m	m	m	m	39.0	(1.5)	0.3	(0.1)	39.3	(1.7)	0.2	(0.1)	46.4	(1.2)	0.1	(0.1)	47.8	(1.8)	0.1	(0.1)
Croatia	17.0	(0.9)	5.1	(0.5)	18.5	(1.1)	3.7	(0.6)	17.3	(0.9)	4.6	(0.8)	24.6	(1.2)	3.9	(0.4)	25.4	(1.2)	3.6	(0.4)
Cyprus	m	m	m	m	m	m	m	m	38.0	(0.7)	2.0	(0.3)	42.1	(0.8)	1.6	(0.2)	39.0	(1.0)	1.6	(0.2)
Dominican Republic	m	m	m	m	m	m	m	m	m	m	m	m	85.7	(1.1)	0.0	(0.0)	84.8	(1.1)	0.0	(0.0)
Georgia	m	m	m	m	65.6	(1.3)	0.2	(0.1)	m	m	m	m	50.8	(1.3)	0.9	(0.2)	64.4	(1.2)	0.1	(0.1)
Hong Kong (China)*	8.7	(0.8)	15.9	(0.9)	6.6	(0.7)	16.2	(1.0)	5.6	(0.6)	16.7	(1.0)	9.4	(0.7)	7.4	(0.6)	11.6	(0.8)	7.8	(0.7)
Indonesia	61.6	(3.4)	0.0	(0.0)	65.6	(2.3)	0.0	(0.0)	66.6	(2.2)	0.0	(0.0)	56.0	(1.6)	0.1	(0.1)	60.0	(1.5)	0.1	(0.0)
Jordan	44.3	(1.2)	0.6	(0.2)	45.6	(1.7)	0.5	(0.2)	49.6	(1.5)	0.2	(0.2)	49.8	(1.4)	0.2	(0.1)	40.3	(1.4)	0.7	(0.2)
Kazakhstan	m	m	m	m	55.4	(1.6)	0.3	(0.2)	41.9	(1.8)	0.2	(0.1)	m	m	m	m	60.3	(1.0)	0.4	(0.1)
Kosovo	m	m	m	m	m	m	m	m	m	m	m	m	67.7	(1.1)	0.0	(0.0)	76.5	(0.7)	0.0	(0.0)
Lebanon	m	m	m	m	m	m	m	m	m	m	m	m	62.6	(1.7)	0.4	(0.1)	62.3	(1.6)	0.5	(0.2)
Macao (China)	10.3	(0.5)	5.3	(0.4)	9.6	(0.4)	4.8	(0.5)	8.8	(0.5)	6.7	(0.4)	8.1	(0.4)	9.2	(0.5)	6.0	(0.5)	13.6	(0.6)
Malaysia	m	m	m	m	43.0	(1.5)	0.2	(0.1)	45.5	(1.6)	0.3	(0.1)	m	m	m	m	36.6	(1.3)	0.6	(0.2)
Malta	m	m	m	m	32.5	(0.8)	6.0	(0.6)	m	m	m	m	32.5	(0.8)	7.6	(0.5)	33.5	(0.9)	4.4	(0.4)
Moldova	m	m	m	m	47.3	(1.5)	0.2	(0.1)	m	m	m	m	42.2	(1.1)	0.7	(0.2)	42.6	(1.2)	0.9	(0.2)
Montenegro	50.2	(0.9)	0.3	(0.1)	53.6	(1.0)	0.2	(0.1)	50.7	(0.7)	0.4	(0.1)	51.0	(0.7)	0.5	(0.1)	48.2	(0.7)	0.3	(0.1)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	69.4	(1.8)	0.0	(0.0)
North Macedonia	m	m	m	m	m	m	m	m	m	m	m	m	62.9	(0.8)	0.2	(0.1)	49.5	(0.8)	0.8	(0.2)
Panama	m	m	m	m	65.1	(2.8)	0.2	(0.1)	m	m	m	m	m	m	m	m	71.3	(1.4)	0.1	(0.1)
Peru	m	m	m	m	68.3	(1.7)	0.2	(0.1)	68.5	(2.0)	0.0	(0.1)	58.5	(1.4)	0.1	(0.1)	54.5	(1.4)	0.2	(0.1)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	78.0	(1.5)	0.1	(0.0)
Qatar	79.1	(0.4)	0.3	(0.1)	65.2	(0.6)	1.4	(0.1)	62.6	(0.5)	1.5	(0.1)	49.8	(0.5)	1.7	(0.2)	48.4	(0.5)	2.2	(0.2)
Romania	46.9	(2.4)	0.5	(0.1)	41.4	(2.1)	0.4	(0.1)	37.3	(1.6)	0.9	(0.3)	38.5	(1.8)	0.7	(0.2)	43.9	(2.1)	1.0	(0.3)
Russia	22.2	(1.4)	4.2	(0.5)	22.0	(1.4)	4.4	(0.5)	18.8	(1.1)	4.3	(0.6)	18.2	(1.1)	3.7	(0.4)	21.2	(1.2)	3.1	(0.4)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	62.3	(1.5)	0.1	(0.0)
Serbia	38.5	(1.6)	0.8	(0.2)	34.4	(1.3)	1.0	(0.2)	35.0	(1.8)	1.7	(0.4)	m	m	m	m	38.3	(1.5)	1.6	(0.2)
Singapore	m	m	m	m	11.5	(0.5)	19.9	(0.6)	9.6	(0.5)	22.7	(0.8)	9.6	(0.4)	24.2	(0.6)	9.0	(0.4)	20.7	(0.6)
Chinese Taipei	11.6	(1.0)	14.6	(0.9)	11.1	(0.7)	8.8	(0.9)	9.8	(0.8)	8.3	(0.6)	12.4	(0.8)	15.4	(1.1)	15.1	(0.8)	11.7	(0.9)
Thailand	46.1	(1.2)	0.4	(0.1)	42.8	(1.6)	0.6	(0.3)	33.6	(1.6)	0.9	(0.3)	46.7	(1.5)	0.5	(0.2)	44.5	(1.5)	0.7	(0.2)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	26.4	(1.4)	3.5	(0.5)
United Arab Emirates	m	m	m	m	39.2	(1.2)	2.2	(0.2)	35.2	(1.3)	2.5	(0.3)	41.8	(1.1)	2.8	(0.2)	42.8	(0.9)	2.9	(0.2)
Uruguay	42.1	(1.4)	1.4	(0.2)	42.6	(1.1)	1.5	(0.2)	46.9	(1.3)	1.0	(0.2)	40.8	(1.1)	1.3	(0.2)	43.9	(1.3)	0.7	(0.2)
Viet Nam**	m	m	m	m	m	m	m	m	6.7	(1.1)	8.1	(1.1)	5.9	(0.8)	8.3	(1.2)	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.9 [3/4] **Percentage of low achievers and top performers in science, 2006 through 2018**

	Change between 2006 and 2018 (PISA 2018 - PISA 2006)				Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
	Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)	
	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
OECD Australia	**6.0**	(1.2)	**-5.1**	(1.0)	**6.3**	(1.3)	**-5.1**	(1.1)	**5.2**	(1.4)	**-4.1**	(1.0)	1.2	(0.8)	**-1.7**	(0.7)
Austria	**5.5**	(2.1)	**-3.7**	(1.0)	m	m	m	m	**6.1**	(2.2)	-1.6	(1.0)	1.0	(1.4)	-1.4	(0.8)
Belgium	3.0	(1.5)	**-2.0**	(0.8)	2.0	(1.5)	**-2.1**	(1.0)	2.3	(1.6)	-1.0	(0.9)	0.2	(1.3)	-1.0	(0.6)
Canada	**3.4**	(1.0)	**-3.1**	(1.0)	**3.9**	(1.0)	-0.8	(1.0)	**3.0**	(1.1)	0.0	(1.1)	**2.3**	(0.8)	-1.0	(0.9)
Chile	-4.4	(3.2)	**-0.9**	(0.4)	3.1	(3.0)	-0.1	(0.3)	0.9	(3.6)	0.0	(0.2)	0.5	(1.7)	-0.2	(0.3)
Colombia	**-9.8**	(4.0)	**0.3**	(0.1)	-3.7	(4.3)	**0.3**	(0.1)	-5.8	(4.8)	**0.3**	(0.1)	1.4	(2.2)	0.1	(0.1)
Czech Republic	3.2	(1.9)	**-4.1**	(1.1)	1.5	(1.9)	-0.8	(0.9)	**5.0**	(2.0)	-0.1	(0.9)	-1.9	(1.5)	0.2	(0.7)
Denmark	0.2	(1.6)	-1.3	(0.9)	2.1	(1.4)	-1.2	(0.8)	2.0	(1.7)	-1.2	(0.9)	**2.8**	(1.1)	-1.5	(0.8)
Estonia	1.1	(1.0)	0.7	(1.2)	0.4	(1.1)	1.8	(1.2)	**3.7**	(1.0)	-0.6	(1.3)	0.0	(0.9)	-1.3	(0.9)
Finland	**8.8**	(1.0)	**-8.6**	(1.3)	**6.9**	(1.0)	**-6.4**	(1.4)	**5.2**	(1.2)	**-4.8**	(1.5)	1.4	(1.0)	**-2.1**	(0.9)
France	-0.7	(1.8)	-1.5	(0.9)	1.2	(1.8)	-1.6	(1.0)	1.8	(1.7)	-1.3	(1.1)	-1.6	(1.2)	-1.4	(0.8)
Germany	**4.2**	(1.8)	-1.8	(1.0)	**4.8**	(1.6)	**-2.8**	(1.1)	**7.4**	(1.7)	-2.2	(1.2)	2.6	(1.4)	-0.6	(0.9)
Greece	**7.7**	(2.9)	**-2.1**	(0.4)	**6.5**	(3.1)	**-1.7**	(0.4)	6.2	(3.5)	**-1.2**	(0.5)	-1.0	(2.4)	**-0.8**	(0.4)
Hungary	**9.1**	(1.9)	**-2.2**	(0.8)	**10.0**	(2.2)	-0.7	(0.8)	**6.1**	(2.3)	-1.2	(1.0)	-1.9	(1.4)	0.1	(0.7)
Iceland	**4.4**	(2.0)	**-2.5**	(0.7)	**7.0**	(2.1)	**-3.1**	(0.6)	1.0	(2.4)	-1.4	(0.8)	-0.3	(1.3)	0.1	(0.5)
Ireland	1.5	(1.6)	**-3.6**	(0.9)	1.9	(1.6)	**-2.9**	(1.0)	**5.9**	(1.6)	**-4.9**	(0.9)	1.7	(1.3)	-1.2	(0.7)
Israel	-3.0	(2.2)	0.6	(0.8)	0.0	(2.2)	**1.9**	(0.7)	4.2	(2.6)	0.0	(0.9)	1.7	(2.0)	0.0	(0.7)
Italy	0.6	(2.0)	**-1.9**	(0.5)	**5.2**	(1.9)	**-3.1**	(0.5)	**7.2**	(2.3)	**-3.3**	(0.6)	2.7	(1.5)	**-1.3**	(0.6)
Japan	-1.2	(1.4)	-2.0	(1.6)	0.2	(1.4)	**-3.9**	(1.7)	2.4	(1.4)	**-5.2**	(2.0)	1.2	(1.0)	-2.3	(1.4)
Korea	2.9	(1.5)	1.5	(1.5)	**7.8**	(1.3)	0.2	(1.5)	**7.5**	(1.3)	0.1	(1.7)	-0.2	(1.2)	1.2	(1.2)
Latvia	1.1	(1.8)	-0.4	(0.6)	**3.8**	(1.9)	0.6	(0.6)	**6.1**	(1.9)	-0.7	(0.7)	1.2	(1.1)	-0.1	(0.5)
Lithuania	1.8	(1.7)	-0.5	(0.8)	**5.2**	(1.8)	-0.2	(0.6)	**6.1**	(2.0)	-0.7	(0.6)	-2.5	(1.4)	0.3	(0.6)
Luxembourg	**4.7**	(1.5)	-0.4	(0.7)	3.1	(1.7)	-1.3	(0.8)	**4.5**	(2.0)	**-2.7**	(0.9)	0.9	(0.9)	**-1.5**	(0.6)
Mexico	-4.1	(4.2)	0.0	(0.1)	-0.5	(4.4)	0.1	(0.1)	-0.2	(5.2)	0.1	(0.1)	-1.0	(2.1)	0.1	(0.1)
Netherlands*	**7.1**	(1.7)	-2.5	(1.3)	**6.9**	(2.1)	-2.1	(1.6)	**6.9**	(1.9)	-1.2	(1.5)	1.5	(1.5)	-0.5	(1.0)
New Zealand	**4.3**	(1.2)	**-6.3**	(1.1)	**4.7**	(1.3)	**-6.3**	(1.2)	1.8	(1.4)	-2.0	(1.2)	0.6	(1.2)	-1.5	(0.9)
Norway	-0.2	(1.9)	0.7	(0.8)	**5.1**	(1.8)	0.4	(0.9)	1.2	(2.0)	-0.8	(0.9)	2.1	(1.3)	-1.2	(0.7)
Poland	**-3.1**	(1.3)	**2.5**	(1.2)	0.7	(1.4)	1.7	(1.2)	**4.8**	(1.4)	-1.5	(1.6)	**-2.4**	(1.2)	1.9	(1.0)
Portugal*	**-4.9**	(1.9)	**2.5**	(0.7)	3.1	(1.7)	1.5	(0.9)	0.6	(2.1)	1.1	(0.9)	2.2	(1.4)	**-1.8**	(0.8)
Slovak Republic	**9.1**	(2.3)	**-2.1**	(0.7)	**10.0**	(2.4)	**-2.5**	(0.7)	2.4	(3.0)	-1.2	(0.8)	-1.4	(1.5)	0.1	(0.5)
Slovenia	0.7	(1.1)	**-5.6**	(0.9)	-0.2	(1.1)	**-2.5**	(1.0)	1.7	(1.3)	**-2.2**	(1.1)	-0.4	(0.8)	**-3.3**	(0.8)
Spain	1.6	(1.6)	-0.7	(0.5)	3.0	(1.7)	0.2	(0.4)	**5.6**	(1.8)	-0.6	(0.5)	**3.0**	(1.0)	-0.8	(0.5)
Sweden	2.6	(1.5)	0.4	(0.9)	-0.1	(1.7)	0.2	(0.9)	-3.2	(1.8)	**2.0**	(0.9)	-2.6	(1.6)	-0.2	(0.9)
Switzerland	**4.2**	(1.8)	**-2.7**	(1.2)	**6.2**	(1.8)	**-2.9**	(1.2)	**7.4**	(2.0)	-1.5	(1.2)	1.8	(1.5)	**-2.0**	(1.0)
Turkey	**-21.4**	(2.7)	**1.5**	(0.6)	-4.8	(2.7)	**1.3**	(0.6)	-1.2	(3.1)	0.7	(0.6)	**-19.3**	(2.4)	**2.2**	(0.5)
United Kingdom	0.7	(1.5)	**-4.1**	(1.0)	2.4	(1.6)	-1.7	(1.1)	2.5	(1.9)	-1.5	(1.2)	0.0	(1.2)	-1.2	(0.9)
United States*	**-5.7**	(2.1)	0.0	(1.2)	0.5	(1.8)	0.0	(1.4)	0.5	(2.0)	1.7	(1.3)	-1.7	(1.6)	0.6	(1.0)
OECD average-36b	1.0	(1.1)	**-1.6**	(0.3)	**3.2**	(1.2)	**-1.3**	(0.4)	**3.3**	(1.5)	**-1.2**	(0.4)	-0.1	(0.3)	**-0.7**	(0.1)
OECD average-37	1.1	(1.1)	**-1.6**	(0.3)	m	m	m	m	**3.4**	(1.5)	**-1.2**	(0.4)	-0.1	(0.3)	**-0.7**	(0.1)

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.9 [4/4] **Percentage of low achievers and top performers in science, 2006 through 2018**

	Change between 2006 and 2018 (PISA 2018 - PISA 2006)				Change between 2009 and 2018 (PISA 2018 - PISA 2009)				Change between 2012 and 2018 (PISA 2018 - PISA 2012)				Change between 2015 and 2018 (PISA 2018 - PISA 2015)			
	Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)		Below Level 2 (less than 409.54 score points)		Level 5 or above (at or above 633.33 score points)	
	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.	% dif.	S.E.
Partners																
Albania	m	m	m	m	**-10.3**	(4.4)	0.1	(0.1)	-6.1	(4.9)	-0.2	(0.2)	**5.3**	(2.2)	-0.1	(0.2)
Argentina	-2.8	(3.5)	0.0	(0.2)	1.1	(3.2)	-0.2	(0.2)	2.6	(3.8)	0.2	(0.2)	m	m	m	m
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	**-5.6**	(2.3)	0.2	(0.3)	1.2	(2.4)	0.2	(0.2)	0.2	(2.6)	**0.5**	(0.2)	-1.2	(1.5)	0.1	(0.2)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	3.9	(3.5)	**-1.5**	(0.7)	**7.7**	(3.7)	-1.1	(0.6)	**9.6**	(3.7)	**-1.6**	(0.7)	**8.6**	(2.5)	**-1.4**	(0.6)
Costa Rica	m	m	m	m	**8.8**	(4.4)	-0.2	(0.1)	8.5	(5.3)	-0.1	(0.1)	1.5	(2.2)	0.0	(0.1)
Croatia	**8.4**	(2.0)	**-1.5**	(0.6)	**6.9**	(2.2)	-0.1	(0.7)	**8.1**	(2.4)	-0.9	(0.9)	0.7	(1.7)	-0.3	(0.5)
Cyprus	m	m	m	m	m	m	m	m	0.9	(2.8)	-0.3	(0.4)	**-3.2**	(1.3)	0.1	(0.3)
Dominican Republic	m	m	m	m	m	m	m	m	m	m	m	m	-0.9	(1.6)	0.0	(0.0)
Georgia	m	m	m	m	-1.2	(3.2)	-0.1	(0.1)	m	m	m	m	**13.6**	(1.8)	**-0.8**	(0.2)
Hong Kong (China)*	**2.8**	(1.3)	**-8.1**	(1.3)	**4.9**	(1.3)	**-8.4**	(1.3)	**6.0**	(1.3)	**-8.9**	(1.4)	**2.1**	(1.1)	0.5	(1.0)
Indonesia	-1.6	(5.5)	0.0	(0.1)	-5.6	(5.1)	0.1	(0.0)	-6.6	(6.1)	0.1	(0.0)	4.1	(2.3)	0.0	(0.1)
Jordan	-4.0	(2.9)	0.0	(0.3)	-5.3	(3.3)	0.2	(0.2)	**-9.2**	(3.7)	0.4	(0.2)	**-9.5**	(2.0)	**0.5**	(0.2)
Kazakhstan	m	m	m	m	4.9	(4.3)	0.1	(0.2)	**18.4**	(5.3)	0.2	(0.1)	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m	m	m	**8.8**	(1.3)	0.0	(0.0)
Lebanon	m	m	m	m	m	m	m	m	m	m	m	m	-0.4	(2.4)	0.1	(0.2)
Macao (China)	**-4.3**	(0.8)	**8.3**	(1.3)	**-3.6**	(0.7)	**8.9**	(1.3)	**-2.8**	(0.8)	**7.0**	(1.5)	**-2.1**	(0.7)	**4.5**	(0.8)
Malaysia	m	m	m	m	-6.4	(3.8)	0.4	(0.2)	**-8.9**	(4.5)	0.3	(0.2)	m	m	m	m
Malta	m	m	m	m	1.0	(1.5)	**-1.6**	(0.7)	m	m	m	m	1.0	(1.2)	**-3.2**	(0.6)
Moldova	m	m	m	m	-4.7	(3.7)	**0.6**	(0.2)	m	m	m	m	0.4	(1.7)	0.1	(0.3)
Montenegro	-2.0	(3.3)	0.1	(0.2)	-5.3	(3.6)	0.1	(0.2)	-2.5	(4.3)	-0.1	(0.2)	**-2.8**	(1.1)	-0.1	(0.2)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	m	m	m	m	m	m	**-13.5**	(1.2)	**0.6**	(0.2)
Panama	m	m	m	m	6.2	(3.7)	-0.1	(0.1)	m	m	m	m	m	m	m	m
Peru	m	m	m	m	**-13.8**	(3.8)	0.0	(0.2)	**-14.0**	(4.6)	0.2	(0.1)	**-4.0**	(2.0)	0.1	(0.1)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	**-30.7**	(1.9)	**1.9**	(0.2)	**-16.8**	(2.1)	**0.8**	(0.2)	**-14.3**	(2.5)	**0.8**	(0.2)	-1.4	(0.8)	**0.5**	(0.2)
Romania	-3.0	(4.1)	0.5	(0.3)	2.6	(4.0)	**0.6**	(0.3)	6.6	(4.3)	0.1	(0.4)	5.4	(2.9)	0.3	(0.3)
Russia	-1.0	(2.5)	-1.1	(0.6)	-0.8	(2.5)	-1.3	(0.7)	2.5	(2.7)	-1.2	(0.7)	3.1	(1.7)	-0.7	(0.5)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	-0.2	(2.8)	**0.8**	(0.3)	3.9	(2.7)	**0.5**	(0.3)	3.3	(3.3)	-0.1	(0.4)	m	m	m	m
Singapore	m	m	m	m	**-2.5**	(0.7)	0.8	(1.9)	-0.6	(0.8)	-2.0	(2.3)	-0.6	(0.6)	**-3.4**	(0.9)
Chinese Taipei	**3.5**	(1.4)	**-3.0**	(1.4)	**4.1**	(1.2)	**2.8**	(1.4)	**5.3**	(1.3)	**3.3**	(1.3)	**2.7**	(1.1)	**-3.7**	(1.4)
Thailand	-1.6	(3.3)	0.3	(0.2)	1.7	(3.6)	0.1	(0.3)	**10.9**	(4.2)	-0.2	(0.3)	-2.3	(2.2)	0.3	(0.2)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	m	m	3.6	(2.6)	**0.7**	(0.3)	**7.6**	(3.1)	0.3	(0.4)	1.0	(1.5)	0.1	(0.3)
Uruguay	1.7	(3.1)	**-0.7**	(0.3)	1.3	(3.1)	**-0.8**	(0.3)	-3.0	(3.7)	-0.3	(0.3)	3.1	(1.8)	**-0.6**	(0.3)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.10 [1/4] **Mean reading performance, 2000 through 2018**

		Reading performance, by PISA cycle													
		PISA 2000		PISA 2003		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
OECD	Australia	528	(3.5)	525	(2.1)	513	(2.1)	515	(2.3)	512	(1.6)	503	(1.7)	503	(1.6)
	Austria	492	(2.7)	491	(3.8)	490	(4.1)	m	m	490	(2.8)	485	(2.8)	484	(2.7)
	Belgium	507	(3.6)	507	(2.6)	501	(3.0)	506	(2.3)	509	(2.3)	499	(2.4)	493	(2.3)
	Canada	534	(1.6)	528	(1.7)	527	(2.4)	524	(1.5)	523	(1.9)	527	(2.3)	520	(1.8)
	Chile	410	(3.6)	m	m	442	(5.0)	449	(3.1)	441	(2.9)	459	(2.6)	452	(2.6)
	Colombia	m	m	m	m	385	(5.1)	413	(3.7)	403	(3.4)	425	(2.9)	412	(3.3)
	Czech Republic	492	(2.4)	489	(3.5)	483	(4.2)	478	(2.9)	493	(2.9)	487	(2.6)	490	(2.5)
	Denmark	497	(2.4)	492	(2.8)	494	(3.2)	495	(2.1)	496	(2.6)	500	(2.5)	501	(1.8)
	Estonia	m	m	m	m	501	(2.9)	501	(2.6)	516	(2.0)	519	(2.2)	523	(1.8)
	Finland	546	(2.6)	543	(1.6)	547	(2.1)	536	(2.3)	524	(2.4)	526	(2.5)	520	(2.3)
	France	505	(2.7)	496	(2.7)	488	(4.1)	496	(3.4)	505	(2.8)	499	(2.5)	493	(2.3)
	Germany	484	(2.5)	491	(3.4)	495	(4.4)	497	(2.7)	508	(2.8)	509	(3.0)	498	(3.0)
	Greece	474	(5.0)	472	(4.1)	460	(4.0)	483	(4.3)	477	(3.3)	467	(4.3)	457	(3.6)
	Hungary	480	(4.0)	482	(2.5)	482	(3.3)	494	(3.2)	488	(3.2)	470	(2.7)	476	(2.3)
	Iceland	507	(1.5)	492	(1.6)	484	(1.9)	500	(1.4)	483	(1.8)	482	(2.0)	474	(1.7)
	Ireland	527	(3.2)	515	(2.6)	517	(3.5)	496	(3.0)	523	(2.6)	521	(2.5)	518	(2.2)
	Israel	452	(8.5)	m	m	439	(4.6)	474	(3.6)	486	(5.0)	479	(3.8)	470	(3.7)
	Italy	487	(2.9)	476	(3.0)	469	(2.4)	486	(1.6)	490	(2.0)	485	(2.7)	476	(2.4)
	Japan	522	(5.2)	498	(3.9)	498	(3.6)	520	(3.5)	538	(3.7)	516	(3.2)	504	(2.7)
	Korea	525	(2.4)	534	(3.1)	556	(3.8)	539	(3.5)	536	(3.9)	517	(3.5)	514	(2.9)
	Latvia	458	(5.3)	491	(3.7)	479	(3.7)	484	(3.0)	489	(2.4)	488	(1.8)	479	(1.6)
	Lithuania	m	m	m	m	470	(3.0)	468	(2.4)	477	(2.5)	472	(2.7)	476	(1.5)
	Luxembourg	m	m	479	(1.5)	479	(1.3)	472	(1.3)	488	(1.5)	481	(1.4)	470	(1.1)
	Mexico	422	(3.3)	400	(4.1)	410	(3.1)	425	(2.0)	424	(1.5)	423	(2.6)	420	(2.7)
	Netherlands*	m	m	513	(2.9)	507	(2.9)	508	(5.1)	511	(3.5)	503	(2.4)	485	(2.7)
	New Zealand	529	(2.8)	522	(2.5)	521	(3.0)	521	(2.4)	512	(2.4)	509	(2.4)	506	(2.0)
	Norway	505	(2.8)	500	(2.8)	484	(3.2)	503	(2.6)	504	(3.2)	513	(2.5)	499	(2.2)
	Poland	479	(4.5)	497	(2.9)	508	(2.8)	500	(2.6)	518	(3.1)	506	(2.5)	512	(2.7)
	Portugal*	470	(4.5)	478	(3.7)	472	(3.6)	489	(3.1)	488	(3.8)	498	(2.7)	492	(2.4)
	Slovak Republic	m	m	469	(3.1)	466	(3.1)	477	(2.5)	463	(4.2)	453	(2.8)	458	(2.2)
	Slovenia	m	m	m	m	494	(1.0)	483	(1.0)	481	(1.2)	505	(1.5)	495	(1.2)
	Spain	493	(2.7)	481	(2.6)	461	(2.2)	481	(2.0)	488	(1.9)	496	(2.4)	m	m
	Sweden	516	(2.2)	514	(2.4)	507	(3.4)	497	(2.9)	483	(3.0)	500	(3.5)	506	(3.0)
	Switzerland	494	(4.2)	499	(3.3)	499	(3.1)	501	(2.4)	509	(2.6)	492	(3.0)	484	(3.1)
	Turkey	m	m	441	(5.8)	447	(4.2)	464	(3.5)	475	(4.2)	428	(4.0)	466	(2.2)
	United Kingdom	m	m	m	m	495	(2.3)	494	(2.3)	499	(3.5)	498	(2.8)	504	(2.6)
	United States*	504	(7.0)	495	(3.2)	m	m	500	(3.7)	498	(3.7)	497	(3.4)	505	(3.6)
	OECD average-23	500	(0.7)	497	(0.6)	495	(0.7)	499	(0.6)	501	(0.6)	497	(0.6)	493	(0.5)
	OECD average-27	494	(0.8)	m	m	m	m	m	m	498	(0.6)	495	(0.5)	491	(0.5)
	OECD average-29b	m	m	494	(0.6)	m	m	m	m	498	(0.5)	493	(0.5)	490	(0.5)
	OECD average-35a	m	m	m	m	m	m	491	(0.5)	493	(0.5)	490	(0.5)	487	(0.4)
	OECD average-35b	m	m	m	m	486	(0.6)	m	m	493	(0.5)	490	(0.5)	487	(0.4)
	OECD average-36a	m	m	m	m	m	m	m	m	493	(0.5)	490	(0.5)	487	(0.4)

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.10 [2/4] **Mean reading performance, 2000 through 2018**

		Reading performance, by PISA cycle													
		PISA 2000		PISA 2003		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
Partners	**Albania**	349	(3.3)	m	m	m	m	385	(4.0)	394	(3.2)	405	(4.1)	405	(1.9)
	Argentina	418	(9.9)	m	m	374	(7.2)	398	(4.6)	396	(3.7)	m	m	402	(3.0)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	389	(2.5)
	Belarus	m	m	m	m	m	m	m	m	m	m	m	m	474	(2.4)
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	403	(2.9)
	Brazil	396	(3.1)	403	(4.6)	393	(3.7)	412	(2.7)	407	(2.0)	407	(2.8)	413	(2.1)
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	408	(0.9)
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	555	(2.7)
	Bulgaria	430	(4.9)	m	m	402	(6.9)	429	(6.7)	436	(6.0)	432	(5.0)	420	(3.9)
	Costa Rica	m	m	m	m	m	m	443	(3.2)	441	(3.5)	427	(2.6)	426	(3.4)
	Croatia	m	m	m	m	477	(2.8)	476	(2.9)	485	(3.3)	487	(2.7)	479	(2.7)
	Cyprus	m	m	m	m	m	m	m	m	449	(1.2)	443	(1.7)	424	(1.4)
	Dominican Republic	m	m	m	m	m	m	m	m	m	m	358	(3.1)	342	(2.9)
	Georgia	m	m	m	m	m	m	374	(2.9)	m	m	401	(3.0)	380	(2.2)
	Hong Kong (China)*	525	(2.9)	510	(3.7)	536	(2.4)	533	(2.1)	545	(2.8)	527	(2.7)	524	(2.7)
	Indonesia	371	(4.0)	382	(3.4)	393	(5.9)	402	(3.7)	396	(4.2)	397	(2.9)	371	(2.6)
	Jordan	m	m	m	m	401	(3.3)	405	(3.3)	399	(3.6)	408	(2.9)	419	(2.9)
	Kazakhstan	m	m	m	m	m	m	390	(3.1)	393	(2.7)	m	m	387	(1.5)
	Kosovo	m	m	m	m	m	m	m	m	m	m	347	(1.6)	353	(1.1)
	Lebanon	m	m	m	m	m	m	m	m	m	m	347	(4.4)	353	(4.3)
	Macao (China)	m	m	498	(2.2)	492	(1.1)	487	(0.9)	509	(0.9)	509	(1.3)	525	(1.2)
	Malaysia	m	m	m	m	m	m	414	(2.9)	398	(3.3)	m	m	415	(2.9)
	Malta	m	m	m	m	m	m	442	(1.6)	m	m	447	(1.8)	448	(1.7)
	Moldova	m	m	m	m	m	m	388	(2.8)	m	m	416	(2.5)	424	(2.4)
	Montenegro	m	m	m	m	392	(1.2)	408	(1.7)	422	(1.2)	427	(1.6)	421	(1.1)
	Morocco	m	m	m	m	m	m	m	m	m	m	m	m	359	(3.1)
	North Macedonia	373	(1.9)	m	m	m	m	m	m	m	m	352	(1.4)	393	(1.1)
	Panama	m	m	m	m	m	m	371	(6.5)	m	m	m	m	377	(3.0)
	Peru	327	(4.4)	m	m	m	m	370	(4.0)	384	(4.3)	398	(2.9)	401	(3.0)
	Philippines	m	m	m	m	m	m	m	m	m	m	m	m	340	(3.3)
	Qatar	m	m	m	m	312	(1.2)	372	(0.8)	388	(0.8)	402	(1.0)	407	(0.8)
	Romania	m	m	m	m	396	(4.7)	424	(4.1)	438	(4.0)	434	(4.1)	428	(5.1)
	Russia	462	(4.2)	442	(3.9)	440	(4.3)	459	(3.3)	475	(3.0)	495	(3.1)	479	(3.1)
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	399	(3.0)
	Serbia	m	m	m	m	401	(3.5)	442	(2.4)	446	(3.4)	m	m	439	(3.3)
	Singapore	m	m	m	m	m	m	526	(1.1)	542	(1.4)	535	(1.6)	549	(1.6)
	Chinese Taipei	m	m	m	m	496	(3.4)	495	(2.6)	523	(3.0)	497	(2.5)	503	(2.8)
	Thailand	431	(3.2)	420	(2.8)	417	(2.6)	421	(2.6)	441	(3.1)	409	(3.3)	393	(3.2)
	Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	466	(3.5)
	United Arab Emirates	m	m	m	m	m	m	431	(2.9)	442	(2.5)	434	(2.9)	432	(2.3)
	Uruguay	m	m	434	(3.4)	413	(3.4)	426	(2.6)	411	(3.2)	437	(2.5)	427	(2.8)
	Viet Nam**	m	m	m	m	m	m	m	m	508	(4.4)	487	(3.7)	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.10 [3/4] **Mean reading performance, 2000 through 2018**

	Change in reading performance between PISA 2018 and:												Average 3-year trend in reading performance[1] across PISA assessments (since 2000 or earliest assessment available)			Curvilinear trend in reading performance across PISA assessments[2]			
	PISA 2000 (PISA 2018 - PISA 2000)		PISA 2003 (PISA 2018 - PISA 2003)		PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
OECD																			
Australia	**-26**	(5.6)	**-23**	(8.2)	-10	(5.9)	**-12**	(4.5)	**-9**	(4.4)	0	(4.6)	**-4.4**	(1.3)	0.001	-1.0	(0.5)	0.0	(0.0)
Austria	-8	(5.6)	-6	(9.0)	-6	(7.2)	m	m	-5	(5.4)	0	(5.5)	-1.3	(2.0)	0.512	-0.8	(0.8)	0.0	(0.0)
Belgium	**-14**	(5.9)	-14	(8.5)	-8	(6.5)	**-13**	(4.8)	**-16**	(4.9)	-6	(5.2)	-1.8	(1.3)	0.172	**-1.8**	(0.6)	**-0.1**	(0.0)
Canada	**-14**	(4.7)	-8	(8.2)	-7	(6.1)	-4	(4.2)	-3	(4.6)	-7	(4.9)	-1.7	(1.3)	0.173	0.0	(0.4)	0.0	(0.0)
Chile	**43**	(6.0)	m	m	10	(7.7)	3	(5.4)	**11**	(5.4)	-6	(5.4)	**7.1**	(1.4)	0.000	-1.3	(0.8)	**-0.2**	(0.0)
Colombia	m	m	m	m	**27**	(8.0)	-1	(6.1)	9	(6.0)	**-13**	(5.9)	**6.6**	(1.8)	0.000	-2.5	(1.3)	**-0.4**	(0.1)
Czech Republic	-1	(5.3)	2	(8.9)	8	(7.2)	**12**	(5.2)	-3	(5.4)	3	(5.4)	0.1	(1.3)	0.925	**1.7**	(0.7)	**0.1**	(0.0)
Denmark	4	(5.0)	9	(8.5)	7	(6.4)	6	(4.5)	5	(4.9)	1	(5.0)	1.1	(1.3)	0.398	**1.3**	(0.5)	0.1	(0.0)
Estonia	m	m	m	m	**22**	(6.3)	**22**	(4.8)	7	(4.6)	4	(4.9)	**6.3**	(1.4)	0.000	**1.6**	(0.8)	0.0	(0.1)
Finland	**-26**	(5.3)	**-23**	(8.3)	**-27**	(6.1)	**-16**	(4.8)	-4	(5.0)	-6	(5.2)	**-4.9**	(1.3)	0.000	**-2.2**	(0.5)	0.0	(0.0)
France	**-12**	(5.4)	-4	(8.5)	5	(7.0)	-3	(5.4)	**-13**	(5.2)	-7	(5.2)	-0.4	(1.3)	0.745	0.5	(0.7)	0.0	(0.0)
Germany	**14**	(5.6)	7	(9.0)	3	(7.5)	1	(5.3)	-9	(5.6)	-11	(5.8)	**3.3**	(1.3)	0.015	-0.9	(0.7)	**-0.1**	(0.0)
Greece	**-16**	(7.4)	-15	(9.5)	-2	(7.5)	**-25**	(6.6)	**-20**	(6.1)	-10	(6.9)	-1.5	(1.5)	0.293	**-2.5**	(0.9)	**-0.1**	(0.1)
Hungary	-4	(6.1)	-6	(8.5)	-6	(6.6)	**-18**	(5.2)	**-12**	(5.4)	6	(5.3)	-1.1	(1.4)	0.416	**-3.0**	(0.7)	**-0.1**	(0.0)
Iceland	**-33**	(4.6)	**-18**	(8.1)	-10	(5.9)	**-26**	(4.2)	-9	(4.5)	-8	(4.7)	**-4.4**	(1.3)	0.000	**-1.4**	(0.4)	0.0	(0.0)
Ireland	-9	(5.6)	3	(8.5)	1	(6.7)	**22**	(5.1)	-5	(5.1)	-3	(5.2)	-0.3	(1.3)	0.824	**2.8**	(0.6)	**0.2**	(0.0)
Israel	18	(10.1)	m	m	**32**	(7.9)	-4	(6.3)	**-15**	(7.2)	-9	(6.6)	**6.1**	(1.9)	0.001	-1.0	(1.2)	**-0.2**	(0.1)
Italy	**-11**	(5.5)	1	(8.7)	8	(6.3)	**-10**	(4.6)	**-13**	(4.9)	-8	(5.3)	0.2	(1.3)	0.862	0.1	(0.5)	0.0	(0.0)
Japan	**-18**	(7.1)	6	(9.1)	6	(6.9)	**-16**	(5.6)	**-34**	(5.9)	**-12**	(5.7)	0.8	(1.4)	0.594	-1.1	(0.8)	-0.1	(0.0)
Korea	-11	(5.6)	**-20**	(8.9)	**-42**	(7.1)	**-25**	(5.7)	**-22**	(6.2)	-3	(6.0)	**-3.1**	(1.3)	0.021	**-6.7**	(0.7)	**-0.3**	(0.0)
Latvia	**21**	(6.8)	-12	(8.7)	-1	(6.6)	-5	(4.9)	**-10**	(4.7)	**-9**	(4.6)	2.3	(1.4)	0.094	**-3.0**	(0.7)	**-0.2**	(0.0)
Lithuania	m	m	m	m	6	(6.2)	7	(4.5)	-1	(4.7)	3	(5.0)	1.6	(1.4)	0.268	0.2	(0.8)	0.0	(0.1)
Luxembourg	m	m	-9	(8.0)	-9	(5.5)	-2	(3.9)	**-18**	(4.2)	**-11**	(4.3)	-0.7	(1.3)	0.576	**-1.8**	(0.3)	**-0.1**	(0.0)
Mexico	-1	(5.9)	**21**	(9.2)	10	(6.7)	-5	(4.9)	-3	(4.9)	-3	(5.4)	2.0	(1.3)	0.134	0.9	(0.6)	0.0	(0.0)
Netherlands*	m	m	**-28**	(8.7)	**-22**	(6.6)	**-24**	(6.8)	**-26**	(5.7)	**-18**	(5.3)	**-4.3**	(1.4)	0.002	**-4.4**	(1.0)	**-0.2**	(0.1)
New Zealand	**-23**	(5.3)	-16	(8.4)	**-15**	(6.4)	**-15**	(4.7)	-6	(4.9)	-4	(5.0)	**-3.7**	(1.3)	0.005	**-1.5**	(0.5)	0.0	(0.0)
Norway	-6	(5.4)	0	(8.5)	**15**	(6.5)	-4	(4.9)	-4	(5.4)	**-14**	(5.1)	1.0	(1.3)	0.429	**1.4**	(0.6)	0.1	(0.0)
Poland	**33**	(6.6)	15	(8.7)	4	(6.5)	**11**	(5.1)	-6	(5.6)	6	(5.4)	**4.5**	(1.4)	0.001	**-1.4**	(0.7)	**-0.2**	(0.0)
Portugal*	**22**	(6.5)	14	(9.0)	**19**	(6.8)	2	(5.3)	4	(5.8)	-6	(5.3)	**4.3**	(1.4)	0.002	0.8	(0.7)	0.0	(0.0)
Slovak Republic	m	m	-11	(8.7)	-8	(6.5)	**-19**	(4.9)	-5	(6.0)	5	(5.3)	**-3.2**	(1.4)	0.019	**-2.4**	(0.8)	-0.1	(0.1)
Slovenia	m	m	m	m	1	(5.5)	**12**	(3.9)	**14**	(4.1)	**-10**	(4.4)	2.4	(1.3)	0.062	**3.5**	(0.5)	**0.2**	(0.0)
Spain	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Sweden	-11	(5.5)	-8	(8.7)	-2	(7.0)	8	(5.5)	**22**	(5.7)	6	(6.1)	**-3.0**	(1.3)	0.024	**2.6**	(0.7)	**0.2**	(0.0)
Switzerland	-10	(6.6)	-15	(9.0)	**-15**	(6.8)	**-17**	(5.3)	**-25**	(5.5)	-8	(5.9)	-1.3	(1.4)	0.339	**-3.7**	(0.7)	**-0.2**	(0.0)
Turkey	m	m	**25**	(9.9)	**18**	(7.1)	1	(5.4)	-10	(6.0)	**37**	(6.0)	2.2	(1.6)	0.160	**-2.3**	(1.0)	**-0.2**	(0.1)
United Kingdom	m	m	m	m	9	(6.3)	**10**	(4.9)	5	(5.7)	6	(5.5)	2.1	(1.4)	0.137	1.5	(1.0)	0.1	(0.1)
United States*	1	(8.9)	10	(9.1)	m	m	6	(6.2)	8	(6.4)	8	(6.3)	0.2	(1.6)	0.916	1.7	(1.0)	0.1	(0.1)
OECD average-23	-7	(4.1)	-5	(7.8)	-3	(5.3)	-7	(3.6)	**-9**	(3.8)	-4	(3.8)	-0.5	(1.2)	0.672	**-0.8**	(0.1)	**-0.03**	(0.01)
OECD average-27	-4	(4.1)	m	m	m	m	m	m	-7	(3.8)	-4	(3.8)	0.0	(1.2)	0.997	**-0.7**	(0.1)	**-0.04**	(0.01)
OECD average-29b	m	m	-4	(7.8)	m	m	m	m	**-9**	(3.8)	-3	(4.0)	-0.7	(1.2)	0.598	**-1.0**	(0.1)	**-0.05**	(0.01)
OECD average-35a	m	m	m	m	m	m	-4	(3.6)	-6	(3.8)	-3	(4.0)	0.4	(1.2)	0.736	**-0.7**	(0.1)	**-0.05**	(0.01)
OECD average-35b	m	m	m	m	1	(5.3)	m	m	-7	(3.8)	-3	(4.0)	0.4	(1.2)	0.765	**-0.8**	(0.1)	**-0.06**	(0.01)
OECD average-36a	m	m	m	m	m	m	m	m	-6	(3.8)	-3	(4.0)	0.4	(1.2)	0.769	**-0.7**	(0.1)	**-0.05**	(0.01)

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.10 [4/4] **Mean reading performance, 2000 through 2018**

	Change in reading performance between PISA 2018 and:												Average 3-year trend in reading performance[1] across PISA assessments (since 2000 or earliest assessment available)			Curvilinear trend in reading performance across PISA assessments[2]			
	PISA 2000 (PISA 2018 - PISA 2000)		PISA 2003 (PISA 2018 - PISA 2003)		PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
Partners	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
Albania	**57**	(5.6)	m	m	m	m	**21**	(5.7)	**11**	(5.3)	0	(6.0)	**10.5**	(1.4)	0.000	1.1	(0.8)	**-0.1**	(0.0)
Argentina	-17	(11.1)	m	m	**28**	(9.4)	3	(6.5)	6	(6.0)	m	m	-1.2	(2.0)	0.553	**4.3**	(1.2)	**0.3**	(0.1)
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	**17**	(5.5)	10	(9.3)	**20**	(6.8)	1	(4.9)	6	(4.8)	6	(5.2)	2.6	(1.3)	0.051	0.9	(0.6)	0.0	(0.0)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	-11	(7.5)	m	m	18	(9.5)	-9	(8.5)	**-16**	(8.1)	-12	(7.5)	0.8	(1.6)	0.598	0.4	(1.3)	0.0	(0.1)
Costa Rica	m	m	m	m	m	m	**-16**	(5.8)	**-14**	(6.2)	-1	(5.8)	**-6.8**	(2.1)	0.001	m	m	m	m
Croatia	m	m	m	m	2	(6.5)	3	(5.3)	-6	(5.7)	-8	(5.5)	1.4	(1.5)	0.332	-1.3	(1.1)	-0.1	(0.1)
Cyprus	m	m	m	m	m	m	m	m	**-25**	(4.2)	**-18**	(4.5)	**-12.2**	(2.1)	0.000	m	m	m	m
Dominican Republic	m	m	m	m	m	m	m	m	m	m	**-16**	(5.7)	**-16.1**	(5.7)	0.005	m	m	m	m
Georgia	m	m	m	m	m	m	5	(5.0)	m	m	**-22**	(5.4)	3.5	(1.9)	0.061	m	m	m	m
Hong Kong (China)*	-1	(5.7)	15	(9.0)	-12	(6.4)	-9	(4.9)	**-20**	(5.4)	-2	(5.5)	1.6	(1.4)	0.251	**-4.0**	(0.7)	**-0.3**	(0.0)
Indonesia	0	(6.2)	-11	(8.9)	**-22**	(8.3)	**-31**	(5.7)	**-25**	(6.2)	**-26**	(5.5)	1.2	(1.4)	0.403	**-6.6**	(0.9)	**-0.4**	(0.1)
Jordan	m	m	m	m	**18**	(6.8)	**14**	(5.7)	**20**	(5.9)	11	(5.7)	**4.0**	(1.6)	0.010	**4.0**	(1.2)	**0.2**	(0.1)
Kazakhstan	m	m	m	m	m	m	-4	(4.9)	-6	(4.8)	m	m	-1.4	(1.6)	0.376	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m	6	(4.4)	5.9	(4.4)	0.175	m	m	m	m
Lebanon	m	m	m	m	m	m	m	m	m	m	7	(7.3)	6.8	(7.3)	0.352	m	m	m	m
Macao (China)	m	m	**27**	(8.2)	**33**	(5.5)	**38**	(3.8)	**16**	(4.0)	**16**	(4.3)	**6.0**	(1.3)	0.000	**5.9**	(0.4)	**0.3**	(0.0)
Malaysia	m	m	m	m	m	m	1	(5.4)	**17**	(5.8)	m	m	2.2	(2.0)	0.260	m	m	m	m
Malta	m	m	m	m	m	m	6	(4.2)	m	m	2	(4.6)	2.3	(1.5)	0.127	m	m	m	m
Moldova	m	m	m	m	m	m	**36**	(5.1)	m	m	8	(5.3)	**13.7**	(1.9)	0.000	m	m	m	m
Montenegro	m	m	m	m	**29**	(5.5)	**14**	(4.1)	-1	(4.1)	-6	(4.4)	**7.7**	(1.3)	0.000	**-2.4**	(0.4)	**-0.4**	(0.0)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	**20**	(4.6)	m	m	m	m	m	m	m	m	**41**	(4.3)	1.1	(4.0)	0.791	m	m	m	m
Panama	m	m	m	m	m	m	6	(8.0)	m	m	m	m	2.1	(2.7)	0.435	m	m	m	m
Peru	**73**	(6.7)	m	m	m	m	**31**	(6.1)	**16**	(6.5)	3	(5.7)	**13.5**	(1.5)	0.000	**2.2**	(0.9)	**-0.1**	(0.1)
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	m	m	m	m	**95**	(5.4)	**35**	(3.7)	**20**	(3.9)	5	(4.1)	**21.9**	(1.3)	0.000	**-3.1**	(0.3)	**-0.9**	(0.0)
Romania	m	m	m	m	**32**	(8.7)	3	(7.5)	-10	(7.5)	-6	(7.6)	**7.2**	(1.9)	0.000	**-5.8**	(1.7)	**-0.7**	(0.1)
Russia	**17**	(6.6)	**36**	(9.2)	**39**	(7.5)	**19**	(5.7)	3	(5.7)	**-16**	(5.9)	**6.8**	(1.4)	0.000	**5.1**	(0.8)	**0.2**	(0.0)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	m	m	m	m	**38**	(7.1)	-3	(5.4)	-7	(6.0)	m	m	**7.7**	(1.6)	0.000	m	m	m	m
Singapore	m	m	m	m	m	m	**24**	(4.0)	7	(4.3)	**14**	(4.5)	**6.4**	(1.4)	0.000	m	m	m	m
Chinese Taipei	m	m	m	m	6	(6.8)	7	(5.2)	**-21**	(5.6)	6	(5.5)	1.5	(1.6)	0.339	**-3.4**	(1.1)	**-0.3**	(0.1)
Thailand	**-38**	(6.1)	**-27**	(8.9)	**-24**	(6.7)	**-28**	(5.5)	**-48**	(5.8)	**-16**	(6.1)	**-4.1**	(1.4)	0.003	**-5.3**	(0.7)	**-0.2**	(0.0)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	m	m	m	m	0	(5.1)	**-10**	(5.1)	-2	(5.4)	-0.7	(2.2)	0.738	m	m	m	m
Uruguay	m	m	-7	(8.9)	**15**	(6.8)	1	(5.2)	**16**	(5.6)	-9	(5.4)	0.6	(1.4)	0.656	3.4	(0.8)	0.2	(0.1)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	m	m	**-21.4**	(6.1)	0.000	m	m	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+. Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.11 [1/4] **Mean mathematics performance, 2003 through 2018**

		Mathematics performance, by PISA cycle											
		PISA 2003		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
OECD	Australia	524	(2.1)	520	(2.2)	514	(2.5)	504	(1.6)	494	(1.6)	491	(1.9)
	Austria	506	(3.3)	505	(3.7)	m	m	506	(2.7)	497	(2.9)	499	(3.0)
	Belgium	529	(2.3)	520	(3.0)	515	(2.3)	515	(2.1)	507	(2.4)	508	(2.3)
	Canada	532	(1.8)	527	(2.0)	527	(1.6)	518	(1.8)	516	(2.3)	512	(2.4)
	Chile	m	m	411	(4.6)	421	(3.1)	423	(3.1)	423	(2.5)	417	(2.4)
	Colombia	m	m	370	(3.8)	381	(3.2)	376	(2.9)	390	(2.3)	391	(3.0)
	Czech Republic	516	(3.5)	510	(3.6)	493	(2.8)	499	(2.9)	492	(2.4)	499	(2.5)
	Denmark	514	(2.7)	513	(2.6)	503	(2.6)	500	(2.3)	511	(2.2)	509	(1.7)
	Estonia	m	m	515	(2.7)	512	(2.6)	521	(2.0)	520	(2.0)	523	(1.7)
	Finland	544	(1.9)	548	(2.3)	541	(2.2)	519	(1.9)	511	(2.3)	507	(2.0)
	France	511	(2.5)	496	(3.2)	497	(3.1)	495	(2.5)	493	(2.1)	495	(2.3)
	Germany	503	(3.3)	504	(3.9)	513	(2.9)	514	(2.9)	506	(2.9)	500	(2.6)
	Greece	445	(3.9)	459	(3.0)	466	(3.9)	453	(2.5)	454	(3.8)	451	(3.1)
	Hungary	490	(2.8)	491	(2.9)	490	(3.5)	477	(3.2)	477	(2.5)	481	(2.3)
	Iceland	515	(1.4)	506	(1.8)	507	(1.4)	493	(1.7)	488	(2.0)	495	(2.0)
	Ireland	503	(2.4)	501	(2.8)	487	(2.5)	501	(2.2)	504	(2.1)	500	(2.2)
	Israel	m	m	442	(4.3)	447	(3.3)	466	(4.7)	470	(3.6)	463	(3.5)
	Italy	466	(3.1)	462	(2.3)	483	(1.9)	485	(2.0)	490	(2.8)	487	(2.8)
	Japan	534	(4.0)	523	(3.3)	529	(3.3)	536	(3.6)	532	(3.0)	527	(2.5)
	Korea	542	(3.2)	547	(3.8)	546	(4.0)	554	(4.6)	524	(3.7)	526	(3.1)
	Latvia	483	(3.7)	486	(3.0)	482	(3.1)	491	(2.8)	482	(1.9)	496	(2.0)
	Lithuania	m	m	486	(2.9)	477	(2.6)	479	(2.6)	478	(2.3)	481	(2.0)
	Luxembourg	493	(1.0)	490	(1.1)	489	(1.2)	490	(1.1)	486	(1.3)	483	(1.1)
	Mexico	385	(3.6)	406	(2.9)	419	(1.8)	413	(1.4)	408	(2.2)	409	(2.5)
	Netherlands*	538	(3.1)	531	(2.6)	526	(4.7)	523	(3.5)	512	(2.2)	519	(2.6)
	New Zealand	523	(2.3)	522	(2.4)	519	(2.3)	500	(2.2)	495	(2.3)	494	(1.7)
	Norway	495	(2.4)	490	(2.6)	498	(2.4)	489	(2.7)	502	(2.2)	501	(2.2)
	Poland	490	(2.5)	495	(2.4)	495	(2.8)	518	(3.6)	504	(2.4)	516	(2.6)
	Portugal*	466	(3.4)	466	(3.1)	487	(2.9)	487	(3.8)	492	(2.5)	492	(2.7)
	Slovak Republic	498	(3.3)	492	(2.8)	497	(3.1)	482	(3.4)	475	(2.7)	486	(2.6)
	Slovenia	m	m	504	(1.0)	501	(1.2)	501	(1.2)	510	(1.3)	509	(1.4)
	Spain	485	(2.4)	480	(2.3)	483	(2.1)	484	(1.9)	486	(2.2)	481	(1.5)
	Sweden	509	(2.6)	502	(2.4)	494	(2.9)	478	(2.3)	494	(3.2)	502	(2.7)
	Switzerland	527	(3.4)	530	(3.2)	534	(3.3)	531	(3.0)	521	(2.9)	515	(2.9)
	Turkey	423	(6.7)	424	(4.9)	445	(4.4)	448	(4.8)	420	(4.1)	454	(2.3)
	United Kingdom	m	m	495	(2.1)	492	(2.4)	494	(3.3)	492	(2.5)	502	(2.6)
	United States*	483	(2.9)	474	(4.0)	487	(3.6)	481	(3.6)	470	(3.2)	478	(3.2)
	OECD average-29a	499	(0.6)	497	(0.5)	499	(0.5)	496	(0.5)	491	(0.5)	494	(0.4)
	OECD average-30	499	(0.6)	497	(0.5)	m	m	496	(0.5)	491	(0.5)	494	(0.4)
	OECD average-36b	m	m	490	(0.5)	492	(0.5)	490	(0.5)	487	(0.4)	489	(0.4)
	OECD average-37	m	m	490	(0.5)	m	m	490	(0.5)	487	(0.4)	489	(0.4)

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.11 [2/4] **Mean mathematics performance, 2003 through 2018**

		Mathematics performance, by PISA cycle											
		PISA 2003		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
Partners	**Albania**	m	m	m	m	377	(4.0)	394	(2.0)	413	(3.4)	437	(2.4)
	Argentina	m	m	381	(6.2)	388	(4.1)	388	(3.5)	m	m	379	(2.8)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	420	(2.8)
	Belarus	m	m	m	m	m	m	m	m	m	m	472	(2.7)
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	406	(3.1)
	Brazil	356	(4.8)	370	(2.9)	386	(2.4)	389	(1.9)	377	(2.9)	384	(2.0)
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	430	(1.2)
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	591	(2.5)
	Bulgaria	m	m	413	(6.1)	428	(5.9)	439	(4.0)	441	(4.0)	436	(3.8)
	Costa Rica	m	m	m	m	409	(3.0)	407	(3.0)	400	(2.5)	402	(3.3)
	Croatia	m	m	467	(2.4)	460	(3.1)	471	(3.5)	464	(2.8)	464	(2.5)
	Cyprus	m	m	m	m	m	m	440	(1.1)	437	(1.7)	451	(1.4)
	Dominican Republic	m	m	m	m	m	m	m	m	328	(2.7)	325	(2.6)
	Georgia	m	m	m	m	379	(2.8)	m	m	404	(2.8)	398	(2.6)
	Hong Kong (China)*	550	(4.5)	547	(2.7)	555	(2.7)	561	(3.2)	548	(3.0)	551	(3.0)
	Indonesia	360	(3.9)	391	(5.6)	371	(3.7)	375	(4.0)	386	(3.1)	379	(3.1)
	Jordan	m	m	384	(3.3)	387	(3.7)	386	(3.1)	380	(2.7)	400	(3.3)
	Kazakhstan	m	m	m	m	405	(3.0)	432	(3.0)	m	m	423	(1.9)
	Kosovo	m	m	m	m	m	m	m	m	362	(1.6)	366	(1.5)
	Lebanon	m	m	m	m	m	m	m	m	396	(3.7)	393	(4.0)
	Macao (China)	527	(2.9)	525	(1.3)	525	(0.9)	538	(1.0)	544	(1.1)	558	(1.5)
	Malaysia	m	m	m	m	404	(2.7)	421	(3.2)	m	m	440	(2.9)
	Malta	m	m	m	m	463	(1.4)	m	m	479	(1.7)	472	(1.9)
	Moldova	m	m	m	m	397	(3.1)	m	m	420	(2.5)	421	(2.4)
	Montenegro	m	m	399	(1.4)	403	(2.0)	410	(1.1)	418	(1.5)	430	(1.2)
	Morocco	m	m	m	m	m	m	m	m	m	m	368	(3.3)
	North Macedonia	m	m	m	m	m	m	m	m	371	(1.3)	394	(1.6)
	Panama	m	m	m	m	360	(5.2)	m	m	m	m	353	(2.7)
	Peru	m	m	m	m	365	(4.0)	368	(3.7)	387	(2.7)	400	(2.6)
	Philippines	m	m	m	m	m	m	m	m	m	m	353	(3.5)
	Qatar	m	m	318	(1.0)	368	(0.7)	376	(0.8)	402	(1.3)	414	(1.2)
	Romania	m	m	415	(4.2)	427	(3.4)	445	(3.8)	444	(3.8)	430	(4.9)
	Russia	468	(4.2)	476	(3.9)	468	(3.3)	482	(3.0)	494	(3.1)	488	(3.0)
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m	373	(3.0)
	Serbia	m	m	435	(3.5)	442	(2.9)	449	(3.4)	m	m	448	(3.2)
	Singapore	m	m	m	m	562	(1.4)	573	(1.3)	564	(1.5)	569	(1.6)
	Chinese Taipei	m	m	549	(4.1)	543	(3.4)	560	(3.3)	542	(3.0)	531	(2.9)
	Thailand	417	(3.0)	417	(2.3)	419	(3.2)	427	(3.4)	415	(3.0)	419	(3.4)
	Ukraine	m	m	m	m	m	m	m	m	m	m	453	(3.6)
	United Arab Emirates	m	m	m	m	421	(2.5)	434	(2.4)	427	(2.4)	435	(2.1)
	Uruguay	422	(3.3)	427	(2.6)	427	(2.6)	409	(2.8)	418	(2.5)	418	(2.6)
	Viet Nam**	m	m	m	m	m	m	511	(4.8)	495	(4.5)	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.11 [3/4] **Mean mathematics performance, 2003 through 2018**

		Change in mathematics performance between PISA 2018 and:										Average 3-year trend in mathematics performance[1] across PISA assessments (since 2003 or earliest assessment available after 2003)			Curvilinear trend in mathematics performance across PISA assessments[2]			
		PISA 2003 (PISA 2018 - PISA 2003)		PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
		Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
OECD	Australia	**-33**	(4.0)	**-29**	(4.3)	**-23**	(4.8)	**-13**	(4.2)	-3	(3.4)	**-7.2**	(1.1)	0.000	**-2.6**	(0.6)	0.0	(0.0)
	Austria	-7	(5.2)	-7	(5.7)	m	m	-7	(5.2)	2	(4.7)	-1.7	(1.2)	0.154	-1.3	(1.0)	0.0	(0.1)
	Belgium	**-21**	(4.3)	**-12**	(4.9)	-7	(4.8)	-6	(4.6)	1	(4.0)	**-4.1**	(1.1)	0.000	-0.2	(0.6)	**0.1**	(0.0)
	Canada	**-20**	(4.1)	**-15**	(4.4)	**-15**	(4.5)	-6	(4.5)	-4	(4.0)	**-4.1**	(1.1)	0.000	**-1.3**	(0.6)	0.0	(0.0)
	Chile	m	m	6	(6.1)	-4	(5.3)	-5	(5.1)	-5	(4.2)	1.4	(1.4)	0.307	**-2.5**	(1.1)	**-0.2**	(0.1)
	Colombia	m	m	**21**	(5.8)	10	(5.7)	**14**	(5.3)	1	(4.4)	**5.1**	(1.3)	0.000	1.5	(1.1)	0.0	(0.1)
	Czech Republic	**-17**	(5.1)	-10	(5.4)	7	(5.2)	1	(5.0)	7	(4.2)	**-3.7**	(1.2)	0.002	**2.1**	(0.8)	**0.2**	(0.1)
	Denmark	-5	(4.3)	-4	(4.5)	6	(4.7)	**9**	(4.4)	-2	(3.6)	-0.9	(1.1)	0.390	**2.1**	(0.6)	**0.2**	(0.0)
	Estonia	m	m	9	(4.5)	**11**	(4.7)	3	(4.3)	4	(3.6)	**2.5**	(1.1)	0.032	1.1	(0.7)	0.0	(0.1)
	Finland	**-37**	(3.9)	**-41**	(4.4)	**-33**	(4.6)	**-11**	(4.3)	-4	(3.8)	**-9.1**	(1.1)	0.000	**-4.1**	(0.6)	**-0.1**	(0.0)
	France	**-15**	(4.4)	0	(5.1)	-1	(5.2)	0	(4.8)	2	(3.9)	**-2.5**	(1.1)	0.025	1.3	(0.7)	**0.1**	(0.0)
	Germany	-3	(5.1)	-4	(5.7)	**-13**	(5.3)	**-13**	(5.1)	-6	(4.6)	-0.1	(1.2)	0.907	**-3.0**	(0.8)	**-0.2**	(0.1)
	Greece	6	(5.7)	-8	(5.3)	**-15**	(6.1)	-2	(5.2)	-2	(5.4)	0.1	(1.3)	0.950	**-3.2**	(0.9)	**-0.2**	(0.1)
	Hungary	-9	(4.6)	**-10**	(4.9)	-9	(5.5)	4	(5.2)	4	(4.1)	**-2.8**	(1.1)	0.013	-0.3	(0.8)	0.0	(0.1)
	Iceland	**-20**	(3.7)	**-10**	(4.1)	**-11**	(4.3)	2	(4.2)	**7**	(3.6)	**-4.7**	(1.0)	0.000	0.2	(0.4)	**0.1**	(0.0)
	Ireland	-3	(4.3)	-2	(4.8)	**12**	(4.9)	-2	(4.6)	-4	(3.8)	0.1	(1.1)	0.897	**1.6**	(0.6)	**0.1**	(0.0)
	Israel	m	m	**21**	(6.4)	**16**	(6.0)	-3	(6.7)	-7	(5.6)	**6.4**	(1.5)	0.000	-1.6	(1.5)	**-0.3**	(0.1)
	Italy	**21**	(5.0)	**25**	(4.8)	4	(4.9)	1	(4.8)	-3	(4.6)	**5.4**	(1.2)	0.000	-0.1	(0.7)	**-0.1**	(0.0)
	Japan	-7	(5.5)	4	(5.2)	-2	(5.5)	-9	(5.5)	-5	(4.5)	0.0	(1.2)	0.998	-0.3	(0.9)	0.0	(0.1)
	Korea	**-16**	(5.3)	**-22**	(5.8)	**-20**	(6.2)	**-28**	(6.5)	2	(5.4)	**-4.1**	(1.2)	0.001	**-5.3**	(1.0)	**-0.3**	(0.1)
	Latvia	**13**	(5.0)	**10**	(4.8)	**14**	(5.1)	6	(4.8)	**14**	(3.6)	1.7	(1.2)	0.162	**1.7**	(0.8)	0.1	(0.1)
	Lithuania	m	m	-5	(4.7)	5	(4.8)	2	(4.7)	3	(3.8)	-0.7	(1.2)	0.532	**2.0**	(0.9)	**0.2**	(0.1)
	Luxembourg	**-10**	(3.2)	-7	(3.5)	-6	(3.9)	-6	(3.7)	-2	(2.9)	-1.7	(1.0)	0.078	**-0.8**	(0.3)	0.0	(0.0)
	Mexico	**24**	(5.2)	3	(5.0)	**-10**	(4.7)	-4	(4.4)	1	(4.1)	**3.4**	(1.2)	0.004	**-4.0**	(0.7)	**-0.3**	(0.0)
	Netherlands*	**-19**	(5.0)	**-11**	(4.9)	-7	(6.5)	-4	(5.5)	7	(4.2)	**-4.2**	(1.2)	0.000	0.0	(0.9)	0.1	(0.1)
	New Zealand	**-29**	(4.0)	**-27**	(4.3)	**-25**	(4.6)	-5	(4.4)	-1	(3.7)	**-7.0**	(1.1)	0.000	**-2.4**	(0.6)	0.0	(0.0)
	Norway	6	(4.3)	**11**	(4.7)	3	(4.8)	**12**	(4.9)	-1	(3.9)	1.5	(1.1)	0.164	**1.7**	(0.7)	0.1	(0.0)
	Poland	**25**	(4.6)	**20**	(4.8)	**21**	(5.2)	-2	(5.6)	**11**	(4.2)	**5.1**	(1.1)	0.000	1.1	(0.8)	0.0	(0.1)
	Portugal*	**26**	(5.2)	**26**	(5.2)	6	(5.3)	5	(5.7)	1	(4.3)	**6.0**	(1.2)	0.000	0.2	(0.9)	**-0.1**	(0.1)
	Slovak Republic	**-12**	(5.1)	-6	(5.0)	**-11**	(5.3)	5	(5.4)	**11**	(4.4)	**-3.6**	(1.2)	0.002	0.1	(0.8)	0.1	(0.1)
	Slovenia	m	m	4	(3.6)	7	(4.0)	**8**	(3.8)	-1	(3.0)	1.8	(0.9)	0.060	**1.8**	(0.5)	**0.1**	(0.0)
	Spain	-4	(4.0)	1	(4.2)	-2	(4.4)	-3	(4.1)	-4	(3.5)	0.0	(1.1)	0.989	-0.1	(0.6)	0.0	(0.0)
	Sweden	-7	(4.6)	0	(4.8)	8	(5.3)	**24**	(4.8)	8	(4.7)	-2.1	(1.1)	0.069	**4.4**	(0.7)	**0.3**	(0.0)
	Switzerland	**-11**	(5.3)	**-14**	(5.3)	**-19**	(5.6)	**-16**	(5.4)	-6	(4.7)	**-2.5**	(1.2)	0.039	**-3.9**	(0.9)	**-0.2**	(0.1)
	Turkey	**30**	(7.6)	**30**	(6.3)	8	(6.1)	6	(6.3)	**33**	(5.3)	**4.1**	(1.5)	0.006	0.4	(1.2)	-0.1	(0.1)
	United Kingdom	m	m	6	(4.6)	9	(5.0)	8	(5.3)	**9**	(4.3)	1.3	(1.1)	0.250	**2.5**	(1.0)	**0.2**	(0.1)
	United States*	-5	(5.2)	4	(6.1)	-9	(6.0)	-3	(5.9)	9	(5.1)	-1.2	(1.2)	0.313	-0.8	(0.9)	0.0	(0.1)
	OECD average-29a	-5	(2.9)	-3	(3.3)	-5	(3.6)	-2	(3.4)	2	(2.4)	-1.3	(1.0)	0.166	**-0.5**	(0.1)	0.0	(0.0)
	OECD average-30	-5	(2.9)	-3	(3.3)	m	m	-2	(3.4)	2	(2.4)	-1.3	(1.0)	0.161	**-0.6**	(0.1)	0.0	(0.0)
	OECD average-36b	m	m	-1	(3.2)	-3	(3.6)	-1	(3.4)	2	(2.4)	-0.6	(0.9)	0.538	**-0.3**	(0.1)	0.0	(0.0)
	OECD average-37	m	m	-1	(3.2)	m	m	-1	(3.4)	2	(2.4)	-0.6	(0.9)	0.516	**-0.3**	(0.1)	0.0	(0.0)

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.11 [4/4] **Mean mathematics performance, 2003 through 2018**

	Change in mathematics performance between PISA 2018 and:										Average 3-year trend in mathematics performance[1] across PISA assessments (since 2003 or earliest assessment available after 2003)			Curvilinear trend in mathematics performance across PISA assessments[2]			
	PISA 2003 (PISA 2018 - PISA 2003)		PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
Partners																	
Albania	m	m	m	m	**60**	(5.9)	**43**	(4.6)	**24**	(4.8)	**19.8**	(1.9)	0.000	m	m	m	m
Argentina	m	m	-2	(7.5)	-9	(6.1)	-9	(5.6)	m	m	-1.0	(1.6)	0.549	m	m	m	m
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	**28**	(5.9)	**14**	(4.8)	-2	(4.7)	-5	(4.4)	6	(4.2)	**4.6**	(1.2)	0.000	**-2.9**	(0.8)	**-0.3**	(0.1)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	m	m	**23**	(7.9)	8	(7.8)	-3	(6.5)	-5	(6.0)	**5.9**	(1.8)	0.001	-2.4	(1.6)	**-0.4**	(0.1)
Costa Rica	m	m	m	m	-7	(5.7)	-5	(5.6)	2	(4.7)	-3.0	(2.0)	0.142	m	m	m	m
Croatia	m	m	-3	(4.7)	4	(5.3)	-7	(5.5)	0	(4.4)	-0.2	(1.2)	0.871	-0.3	(1.1)	0.0	(0.1)
Cyprus	m	m	m	m	m	m	**11**	(3.8)	**14**	(3.2)	**5.7**	(1.9)	0.003	m	m	m	m
Dominican Republic	m	m	m	m	m	m	m	m	-3	(4.4)	-2.6	(4.4)	0.556	m	m	m	m
Georgia	m	m	m	m	**18**	(5.2)	m	m	-6	(4.5)	**7.6**	(2.0)	0.000	m	m	m	m
Hong Kong (China)*	1	(6.1)	4	(5.1)	-3	(5.4)	-10	(5.5)	3	(4.8)	0.4	(1.3)	0.749	-1.3	(0.9)	-0.1	(0.1)
Indonesia	**19**	(5.7)	-12	(7.2)	7	(6.0)	4	(6.1)	-7	(5.0)	2.2	(1.3)	0.086	-1.4	(1.0)	**-0.1**	(0.1)
Jordan	m	m	**16**	(5.7)	**13**	(6.1)	**14**	(5.6)	**20**	(4.8)	2.5	(1.3)	0.057	**3.7**	(1.3)	**0.2**	(0.1)
Kazakhstan	m	m	m	m	**18**	(5.0)	-9	(4.9)	m	m	**4.7**	(1.6)	0.004	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	4	(3.2)	4.6	(3.4)	0.177	m	m	m	m
Lebanon	m	m	m	m	m	m	m	m	-3	(6.0)	-2.8	(6.0)	0.638	m	m	m	m
Macao (China)	**30**	(4.3)	**33**	(3.8)	**32**	(4.0)	**20**	(3.8)	**14**	(3.0)	**6.2**	(1.1)	0.000	**5.0**	(0.4)	**0.2**	(0.0)
Malaysia	m	m	m	m	**36**	(5.3)	**20**	(5.4)	m	m	**12.7**	(1.9)	0.000	m	m	m	m
Malta	m	m	m	m	**9**	(4.3)	m	m	**-7**	(3.5)	**3.9**	(1.6)	0.014	m	m	m	m
Moldova	m	m	m	m	**23**	(5.3)	m	m	1	(4.2)	**9.2**	(2.0)	0.000	m	m	m	m
Montenegro	m	m	**30**	(3.7)	**27**	(4.3)	**20**	(3.7)	**12**	(3.0)	**7.6**	(1.0)	0.000	**4.3**	(0.4)	**0.1**	(0.0)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	m	m	**23**	(3.1)	**23.3**	(3.3)	0.000	m	m	m	m
Panama	m	m	m	m	-7	(6.9)	m	m	m	m	-2.3	(2.3)	0.317	m	m	m	m
Peru	m	m	m	m	**35**	(5.9)	**32**	(5.6)	**13**	(4.4)	**12.2**	(1.9)	0.000	m	m	m	m
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	m	m	**96**	(3.6)	**46**	(3.8)	**38**	(3.6)	**12**	(2.9)	**22.6**	(0.9)	0.000	**1.8**	(0.4)	**-0.5**	(0.0)
Romania	m	m	**15**	(7.2)	3	(6.9)	**-15**	(7.0)	**-14**	(6.6)	**4.7**	(1.7)	0.005	**-5.2**	(1.6)	**-0.6**	(0.1)
Russia	**19**	(5.8)	**12**	(5.8)	**20**	(5.7)	6	(5.4)	-6	(4.9)	**4.7**	(1.3)	0.000	**1.9**	(0.9)	0.0	(0.1)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	m	m	**13**	(5.7)	6	(5.6)	-1	(5.7)	m	m	**3.0**	(1.4)	0.031	m	m	m	m
Singapore	m	m	m	m	7	(4.1)	-4	(3.9)	5	(3.2)	1.1	(1.4)	0.432	m	m	m	m
Chinese Taipei	m	m	**-18**	(5.9)	**-12**	(5.7)	**-29**	(5.5)	**-11**	(4.8)	**-3.8**	(1.4)	0.007	**-5.5**	(1.2)	**-0.4**	(0.1)
Thailand	2	(5.4)	1	(5.2)	0	(5.9)	-8	(5.9)	3	(5.1)	0.3	(1.2)	0.806	-1.0	(0.9)	-0.1	(0.1)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	m	m	**14**	(4.8)	1	(4.7)	7	(4.0)	3.7	(2.0)	0.059	m	m	m	m
Uruguay	-5	(5.1)	-9	(4.9)	-9	(5.1)	8	(5.1)	0	(4.3)	-2.0	(1.2)	0.093	-0.4	(0.8)	0.0	(0.1)
Viet Nam**	m	m	m	m	m	m	m	m	m	m	-17.1	(6.7)	0.011	m	m	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.12 [1/4] **Mean science performance, 2006 through 2018**

		Science performance, by PISA cycle									
		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
OECD	Australia	527	(2.3)	527	(2.5)	521	(1.8)	510	(1.5)	503	(1.8)
	Austria	511	(3.9)	m	m	506	(2.7)	495	(2.4)	490	(2.8)
	Belgium	510	(2.5)	507	(2.5)	505	(2.2)	502	(2.3)	499	(2.2)
	Canada	534	(2.0)	529	(1.6)	525	(1.9)	528	(2.1)	518	(2.2)
	Chile	438	(4.3)	447	(2.9)	445	(2.9)	447	(2.4)	444	(2.4)
	Colombia	388	(3.4)	402	(3.6)	399	(3.1)	416	(2.4)	413	(3.1)
	Czech Republic	513	(3.5)	500	(3.0)	508	(3.0)	493	(2.3)	497	(2.5)
	Denmark	496	(3.1)	499	(2.5)	498	(2.7)	502	(2.4)	493	(1.9)
	Estonia	531	(2.5)	528	(2.7)	541	(1.9)	534	(2.1)	530	(1.9)
	Finland	563	(2.0)	554	(2.3)	545	(2.2)	531	(2.4)	522	(2.5)
	France	495	(3.4)	498	(3.6)	499	(2.6)	495	(2.1)	493	(2.2)
	Germany	516	(3.8)	520	(2.8)	524	(3.0)	509	(2.7)	503	(2.9)
	Greece	473	(3.2)	470	(4.0)	467	(3.1)	455	(3.9)	452	(3.1)
	Hungary	504	(2.7)	503	(3.1)	494	(2.9)	477	(2.4)	481	(2.3)
	Iceland	491	(1.6)	496	(1.4)	478	(2.1)	473	(1.7)	475	(1.8)
	Ireland	508	(3.2)	508	(3.3)	522	(2.5)	503	(2.4)	496	(2.2)
	Israel	454	(3.7)	455	(3.1)	470	(5.0)	467	(3.4)	462	(3.6)
	Italy	475	(2.0)	489	(1.8)	494	(1.9)	481	(2.5)	468	(2.4)
	Japan	531	(3.4)	539	(3.4)	547	(3.6)	538	(3.0)	529	(2.6)
	Korea	522	(3.4)	538	(3.4)	538	(3.7)	516	(3.1)	519	(2.8)
	Latvia	490	(3.0)	494	(3.1)	502	(2.8)	490	(1.6)	487	(1.8)
	Lithuania	488	(2.8)	491	(2.9)	496	(2.6)	475	(2.7)	482	(1.6)
	Luxembourg	486	(1.1)	484	(1.2)	491	(1.3)	483	(1.1)	477	(1.2)
	Mexico	410	(2.7)	416	(1.8)	415	(1.3)	416	(2.1)	419	(2.6)
	Netherlands*	525	(2.7)	522	(5.4)	522	(3.5)	509	(2.3)	503	(2.8)
	New Zealand	530	(2.7)	532	(2.6)	516	(2.1)	513	(2.4)	508	(2.1)
	Norway	487	(3.1)	500	(2.6)	495	(3.1)	498	(2.3)	490	(2.3)
	Poland	498	(2.3)	508	(2.4)	526	(3.1)	501	(2.5)	511	(2.6)
	Portugal*	474	(3.0)	493	(2.9)	489	(3.7)	501	(2.4)	492	(2.8)
	Slovak Republic	488	(2.6)	490	(3.0)	471	(3.6)	461	(2.6)	464	(2.3)
	Slovenia	519	(1.1)	512	(1.1)	514	(1.3)	513	(1.3)	507	(1.3)
	Spain	488	(2.6)	488	(2.1)	496	(1.8)	493	(2.1)	483	(1.6)
	Sweden	503	(2.4)	495	(2.7)	485	(3.0)	493	(3.6)	499	(3.1)
	Switzerland	512	(3.2)	517	(2.8)	515	(2.7)	506	(2.9)	495	(3.0)
	Turkey	424	(3.8)	454	(3.6)	463	(3.9)	425	(3.9)	468	(2.0)
	United Kingdom	515	(2.3)	514	(2.5)	514	(3.4)	509	(2.6)	505	(2.6)
	United States*	489	(4.2)	502	(3.6)	497	(3.8)	496	(3.2)	502	(3.3)
	OECD average-36b	494	(0.5)	498	(0.5)	498	(0.5)	491	(0.4)	489	(0.4)
	OECD average-37	495	(0.5)	m	m	498	(0.5)	491	(0.4)	489	(0.4)

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.12 [2/4] **Mean science performance, 2006 through 2018**

		Science performance, by PISA cycle									
		PISA 2006		PISA 2009		PISA 2012		PISA 2015		PISA 2018	
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.
Partners	**Albania**	m	m	391	(3.9)	397	(2.4)	427	(3.3)	417	(2.0)
	Argentina	391	(6.1)	401	(4.6)	406	(3.9)	m	m	404	(2.9)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	398	(2.4)
	Belarus	m	m	m	m	m	m	m	m	471	(2.4)
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	398	(2.7)
	Brazil	390	(2.8)	405	(2.4)	402	(2.1)	401	(2.3)	404	(2.1)
	Brunei Darussalam	m	m	m	m	m	m	m	m	431	(1.2)
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	590	(2.7)
	Bulgaria	434	(6.1)	439	(5.9)	446	(4.8)	446	(4.4)	424	(3.6)
	Costa Rica	m	m	430	(2.8)	429	(2.9)	420	(2.1)	416	(3.3)
	Croatia	493	(2.4)	486	(2.8)	491	(3.1)	475	(2.5)	472	(2.8)
	Cyprus	m	m	m	m	438	(1.2)	433	(1.4)	439	(1.4)
	Dominican Republic	m	m	m	m	m	m	332	(2.6)	336	(2.5)
	Georgia	m	m	373	(2.9)	m	m	411	(2.4)	383	(2.3)
	Hong Kong (China)*	542	(2.5)	549	(2.8)	555	(2.6)	523	(2.5)	517	(2.5)
	Indonesia	393	(5.7)	383	(3.8)	382	(3.8)	403	(2.6)	396	(2.4)
	Jordan	422	(2.8)	415	(3.5)	409	(3.1)	409	(2.7)	429	(2.9)
	Kazakhstan	m	m	400	(3.1)	425	(3.0)	m	m	397	(1.7)
	Kosovo	m	m	m	m	m	m	378	(1.7)	365	(1.2)
	Lebanon	m	m	m	m	m	m	386	(3.4)	384	(3.5)
	Macao (China)	511	(1.1)	511	(1.0)	521	(0.8)	529	(1.1)	544	(1.5)
	Malaysia	m	m	422	(2.7)	420	(3.0)	m	m	438	(2.7)
	Malta	m	m	461	(1.7)	m	m	465	(1.6)	457	(1.9)
	Moldova	m	m	413	(3.0)	m	m	428	(2.0)	428	(2.3)
	Montenegro	412	(1.1)	401	(2.0)	410	(1.1)	411	(1.0)	415	(1.3)
	Morocco	m	m	m	m	m	m	m	m	377	(3.0)
	North Macedonia	m	m	m	m	m	m	384	(1.2)	413	(1.4)
	Panama	m	m	376	(5.7)	m	m	m	m	365	(2.9)
	Peru	m	m	369	(3.5)	373	(3.6)	397	(2.4)	404	(2.7)
	Philippines	m	m	m	m	m	m	m	m	357	(3.2)
	Qatar	349	(0.9)	379	(0.9)	384	(0.7)	418	(1.0)	419	(0.9)
	Romania	418	(4.2)	428	(3.4)	439	(3.3)	435	(3.2)	426	(4.6)
	Russia	479	(3.7)	478	(3.3)	486	(2.9)	487	(2.9)	478	(2.9)
	Saudi Arabia	m	m	m	m	m	m	m	m	386	(2.8)
	Serbia	436	(3.0)	443	(2.4)	445	(3.4)	m	m	440	(3.0)
	Singapore	m	m	542	(1.4)	551	(1.5)	556	(1.2)	551	(1.5)
	Chinese Taipei	532	(3.6)	520	(2.6)	523	(2.3)	532	(2.7)	516	(2.9)
	Thailand	421	(2.1)	425	(3.0)	444	(2.9)	421	(2.8)	426	(3.2)
	Ukraine	m	m	m	m	m	m	m	m	469	(3.3)
	United Arab Emirates	m	m	438	(2.6)	448	(2.8)	437	(2.4)	434	(2.0)
	Uruguay	428	(2.7)	427	(2.6)	416	(2.8)	435	(2.2)	426	(2.5)
	Viet Nam**	m	m	m	m	528	(4.3)	525	(3.9)	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.12 [3/4] **Mean science performance, 2006 through 2018**

		Change in science performance between PISA 2018 and:								Average 3-year trend in science performance[1] across PISA assessments (since 2006 or earliest assessment available after 2006)			Curvilinear trend in science performance across PISA assessments[2]			
		PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
		Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
OECD	Australia	**-24**	(4.5)	**-24**	(4.7)	**-19**	(4.7)	**-7**	(2.8)	**-6.5**	(1.2)	0.000	**-4.1**	(0.7)	**-0.2**	(0.1)
	Austria	**-21**	(5.9)	m	m	**-16**	(5.6)	-5	(4.0)	**-5.5**	(1.5)	0.000	m	m	m	m
	Belgium	**-12**	(4.8)	-8	(4.9)	-6	(5.1)	-3	(3.5)	**-2.7**	(1.3)	0.035	-0.9	(0.8)	0.0	(0.1)
	Canada	**-16**	(4.6)	**-11**	(4.5)	-7	(4.9)	**-10**	(3.4)	**-3.4**	(1.2)	0.006	-1.3	(0.7)	0.0	(0.1)
	Chile	5	(6.0)	-4	(5.2)	-1	(5.5)	-3	(3.7)	1.1	(1.5)	0.468	-1.6	(1.1)	-0.2	(0.1)
	Colombia	**25**	(5.7)	12	(6.0)	**15**	(5.9)	-2	(4.1)	**6.4**	(1.4)	0.000	1.0	(1.2)	-0.1	(0.1)
	Czech Republic	**-16**	(5.5)	-4	(5.3)	**-12**	(5.6)	4	(3.7)	**-4.0**	(1.4)	0.004	-0.4	(1.0)	0.1	(0.1)
	Denmark	-3	(5.0)	-7	(4.8)	-6	(5.2)	**-9**	(3.4)	-0.4	(1.3)	0.763	**-2.1**	(0.9)	**-0.2**	(0.1)
	Estonia	-1	(4.7)	2	(4.9)	**-11**	(4.8)	-4	(3.2)	0.4	(1.3)	0.771	**-2.0**	(0.8)	**-0.2**	(0.1)
	Finland	**-41**	(4.7)	**-32**	(5.0)	**-24**	(5.2)	**-9**	(3.8)	**-10.7**	(1.3)	0.000	**-4.2**	(0.9)	-0.1	(0.1)
	France	-2	(5.3)	-5	(5.6)	-6	(5.3)	-2	(3.4)	-0.8	(1.4)	0.581	-1.7	(1.0)	-0.1	(0.1)
	Germany	**-13**	(5.9)	**-17**	(5.4)	**-21**	(5.8)	-6	(4.2)	**-3.6**	(1.5)	0.014	**-5.0**	(1.1)	**-0.3**	(0.1)
	Greece	**-22**	(5.7)	**-18**	(6.2)	**-15**	(6.0)	-3	(5.2)	**-5.9**	(1.5)	0.000	**-2.8**	(1.2)	-0.1	(0.1)
	Hungary	**-23**	(5.0)	**-22**	(5.3)	**-13**	(5.5)	4	(3.7)	**-7.1**	(1.3)	0.000	**-2.2**	(1.0)	0.0	(0.1)
	Iceland	**-16**	(4.2)	**-21**	(4.3)	-3	(4.9)	2	(2.9)	**-5.4**	(1.2)	0.000	-1.2	(0.7)	0.0	(0.1)
	Ireland	**-12**	(5.2)	**-12**	(5.3)	**-26**	(5.2)	-6	(3.6)	**-3.0**	(1.4)	0.030	**-5.4**	(0.9)	**-0.4**	(0.1)
	Israel	8	(6.2)	7	(6.0)	-8	(7.3)	-4	(5.2)	2.8	(1.5)	0.066	-1.9	(1.5)	-0.2	(0.1)
	Italy	-7	(4.7)	**-21**	(4.7)	**-26**	(5.1)	**-13**	(3.8)	-2.3	(1.3)	0.063	**-7.4**	(0.8)	**-0.6**	(0.1)
	Japan	-2	(5.5)	-10	(5.6)	**-18**	(6.0)	**-9**	(4.2)	-0.6	(1.4)	0.668	**-5.0**	(1.1)	**-0.4**	(0.1)
	Korea	-3	(5.6)	**-19**	(5.7)	**-19**	(6.1)	3	(4.5)	**-2.9**	(1.4)	0.046	**-5.5**	(1.2)	**-0.4**	(0.1)
	Latvia	-2	(4.9)	-7	(5.0)	**-15**	(5.2)	-3	(2.8)	-0.8	(1.3)	0.519	**-3.7**	(0.9)	**-0.3**	(0.1)
	Lithuania	-6	(4.7)	-9	(4.9)	**-14**	(5.0)	7	(3.5)	**-2.8**	(1.3)	0.032	**-2.7**	(0.8)	**-0.2**	(0.1)
	Luxembourg	**-10**	(3.8)	-7	(4.0)	**-14**	(4.4)	**-6**	(2.2)	-1.9	(1.1)	0.082	**-2.8**	(0.5)	**-0.2**	(0.0)
	Mexico	10	(5.1)	3	(4.8)	4	(4.9)	3	(3.7)	1.9	(1.3)	0.158	0.2	(0.8)	0.0	(0.1)
	Netherlands*	**-21**	(5.3)	**-19**	(7.1)	**-19**	(6.0)	-5	(3.9)	**-5.6**	(1.4)	0.000	**-3.6**	(1.3)	-0.1	(0.1)
	New Zealand	**-22**	(4.9)	**-24**	(4.9)	-7	(5.0)	-5	(3.5)	**-6.2**	(1.3)	0.000	**-1.9**	(0.8)	0.0	(0.1)
	Norway	4	(5.2)	-9	(5.0)	-4	(5.6)	**-8**	(3.6)	0.6	(1.3)	0.657	**-3.0**	(1.0)	**-0.3**	(0.1)
	Poland	**13**	(4.9)	3	(5.1)	**-15**	(5.7)	**10**	(3.9)	2.1	(1.3)	0.109	**-3.4**	(1.0)	**-0.3**	(0.1)
	Portugal*	**17**	(5.4)	-1	(5.4)	2	(6.1)	**-9**	(4.0)	**4.3**	(1.4)	0.002	**-2.4**	(1.2)	**-0.3**	(0.1)
	Slovak Republic	**-24**	(4.9)	**-26**	(5.2)	-7	(5.9)	3	(3.8)	**-7.8**	(1.3)	0.000	-1.5	(1.1)	0.1	(0.1)
	Slovenia	**-12**	(3.9)	-5	(4.0)	-7	(4.4)	**-6**	(2.4)	**-2.2**	(1.1)	0.043	-0.9	(0.4)	0.0	(0.0)
	Spain	-5	(4.6)	-5	(4.4)	**-13**	(4.7)	**-10**	(3.0)	-0.5	(1.2)	0.674	**-3.0**	(0.7)	**-0.2**	(0.1)
	Sweden	-4	(5.2)	4	(5.5)	**15**	(5.9)	6	(5.0)	-1.0	(1.4)	0.463	**4.1**	(1.1)	**0.4**	(0.1)
	Switzerland	**-16**	(5.6)	**-21**	(5.5)	**-20**	(5.7)	**-10**	(4.4)	**-4.4**	(1.4)	0.002	**-5.2**	(1.1)	**-0.3**	(0.1)
	Turkey	**44**	(5.6)	**14**	(5.5)	5	(5.9)	**43**	(4.7)	**6.1**	(1.5)	0.000	0.0	(1.1)	-0.2	(0.1)
	United Kingdom	**-10**	(4.9)	-9	(5.1)	-9	(5.8)	-5	(3.9)	-2.4	(1.3)	0.064	-2.0	(1.0)	-0.1	(0.1)
	United States*	**13**	(6.4)	0	(6.1)	5	(6.4)	6	(4.8)	2.1	(1.6)	0.182	-0.3	(1.3)	-0.1	(0.1)
	OECD average-36b	-6	(3.5)	**-9**	(3.6)	**-9**	(4.1)	-2	(1.6)	-1.9	(1.1)	0.075	**-2.4**	(0.2)	**-0.15**	(0.01)
	OECD average-37	-6	(3.5)	m	m	**-10**	(4.1)	-2	(1.6)	-2.0	(1.1)	0.061	m	m	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.12 [4/4] **Mean science performance, 2006 through 2018**

	Change in science performance between PISA 2018 and:								Average 3-year trend in science performance[1] across PISA assessments (since 2006 or earliest assessment available after 2006)			Curvilinear trend in science performance across PISA assessments[2]			
	PISA 2006 (PISA 2018 - PISA 2006)		PISA 2009 (PISA 2018 - PISA 2009)		PISA 2012 (PISA 2018 - PISA 2012)		PISA 2015 (PISA 2018 - PISA 2015)					Annual rate of change in 2018 (linear term)		Rate of acceleration or deceleration in performance (quadratic term)	
	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Score dif.	S.E.	Coef.	S.E.	p-value	Coef.	S.E.	Coef.	S.E.
Partners															
Albania	m	m	**26**	(5.7)	**19**	(5.1)	**-10**	(4.1)	**10.7**	(2.0)	0.000	m	m	m	m
Argentina	13	(7.6)	3	(6.5)	-2	(6.3)	m	m	3.0	(1.7)	0.076	m	m	m	m
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Brazil	**13**	(4.9)	-2	(4.8)	2	(5.0)	3	(3.4)	2.2	(1.3)	0.087	-1.3	(0.8)	**-0.2**	(0.1)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Bulgaria	-10	(7.9)	-15	(7.8)	**-22**	(7.2)	**-22**	(5.9)	-1.4	(1.9)	0.452	**-6.4**	(1.7)	**-0.5**	(0.1)
Costa Rica	m	m	**-15**	(5.6)	**-14**	(5.9)	-4	(4.2)	**-6.1**	(2.2)	0.006	m	m	m	m
Croatia	**-21**	(5.1)	**-14**	(5.4)	**-19**	(5.8)	-3	(4.0)	**-5.3**	(1.3)	0.000	**-3.1**	(1.1)	-0.1	(0.1)
Cyprus	m	m	m	m	1	(4.4)	**6**	(2.5)	0.7	(2.2)	0.770	m	m	m	m
Dominican Republic	m	m	m	m	m	m	4	(3.9)	4.0	(3.9)	0.306	m	m	m	m
Georgia	m	m	10	(5.2)	m	m	**-28**	(3.7)	**5.6**	(2.0)	0.005	m	m	m	m
Hong Kong (China)*	**-26**	(5.0)	**-32**	(5.2)	**-38**	(5.4)	-7	(3.9)	**-7.7**	(1.3)	0.000	**-8.6**	(1.0)	**-0.5**	(0.1)
Indonesia	3	(7.1)	**14**	(5.7)	**14**	(6.0)	-7	(3.8)	2.5	(1.7)	0.139	**3.6**	(1.3)	0.2	(0.1)
Jordan	7	(5.4)	**14**	(5.8)	**20**	(5.9)	**21**	(4.2)	0.8	(1.4)	0.574	**6.0**	(1.1)	**0.5**	(0.1)
Kazakhstan	m	m	-3	(5.0)	**-28**	(5.3)	m	m	-2.9	(1.6)	0.073	m	m	m	m
Kosovo	m	m	m	m	m	m	**-14**	(2.6)	**-13.6**	(2.6)	0.000	m	m	m	m
Lebanon	m	m	m	m	m	m	-3	(5.1)	-2.8	(5.1)	0.590	m	m	m	m
Macao (China)	**33**	(3.9)	**33**	(4.0)	**23**	(4.4)	**15**	(2.4)	**8.3**	(1.1)	0.000	**5.3**	(0.5)	**0.2**	(0.0)
Malaysia	m	m	**15**	(5.2)	**18**	(5.7)	m	m	**6.6**	(1.9)	0.001	m	m	m	m
Malta	m	m	-5	(4.4)	m	m	**-8**	(2.9)	-1.3	(1.7)	0.433	m	m	m	m
Moldova	m	m	**16**	(5.2)	m	m	0	(3.4)	**6.1**	(2.0)	0.003	m	m	m	m
Montenegro	3	(3.9)	**14**	(4.3)	5	(4.4)	4	(2.2)	1.7	(1.1)	0.144	**2.5**	(0.4)	**0.2**	(0.0)
Morocco	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
North Macedonia	m	m	m	m	m	m	**29**	(2.4)	**28.7**	(2.7)	0.000	m	m	m	m
Panama	m	m	-11	(7.4)	m	m	m	m	-3.8	(2.5)	0.126	m	m	m	m
Peru	m	m	**35**	(5.7)	**31**	(6.0)	8	(3.9)	**12.8**	(2.0)	0.000	m	m	m	m
Philippines	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Qatar	**70**	(3.7)	**40**	(3.8)	**35**	(4.2)	2	(2.0)	**17.9**	(1.1)	0.000	**3.4**	(0.3)	**-0.2**	(0.0)
Romania	7	(7.1)	-2	(6.7)	-13	(6.9)	-9	(5.8)	2.1	(1.7)	0.218	**-4.4**	(1.5)	**-0.4**	(0.1)
Russia	-2	(5.8)	-1	(5.7)	-9	(5.7)	**-9**	(4.4)	0.5	(1.5)	0.755	-2.1	(1.1)	-0.2	(0.1)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Serbia	4	(5.5)	-3	(5.3)	-5	(6.1)	m	m	0.7	(1.3)	0.615	m	m	m	m
Singapore	m	m	**9**	(4.1)	-1	(4.5)	-5	(2.4)	**3.2**	(1.6)	0.039	m	m	m	m
Chinese Taipei	**-17**	(5.7)	-5	(5.3)	-8	(5.5)	**-17**	(4.2)	-2.2	(1.4)	0.118	-1.1	(1.0)	0.0	(0.1)
Thailand	5	(5.2)	1	(5.6)	**-18**	(5.9)	4	(4.5)	0.6	(1.3)	0.677	**-3.8**	(1.1)	**-0.3**	(0.1)
Ukraine	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	m	m	-4	(4.9)	**-15**	(5.3)	-3	(3.5)	-2.5	(2.3)	0.276	m	m	m	m
Uruguay	-2	(5.1)	-1	(5.1)	10	(5.5)	**-10**	(3.6)	0.4	(1.3)	0.746	1.5	(0.9)	0.1	(0.1)
Viet Nam**	m	m	m	m	m	m	m	m	-3.8	(6.2)	0.541	m	m	m	m

1. The average 3-year trend is the average change, per 3-year period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

2. Curvilinear trends in performance are computed only for countries with comparable data in more than four PISA assessments. Curvilinear trends are calculated by a regression of performance over both linear and quadratic (squared) terms of the gap between the year of assessment and 2018. The coefficient for the linear term represents the annual rate of change in 2018, while the coefficient for the quadratic term represents the acceleration or deceleration of the change in performance.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.13 [1/2] **Distribution of reading scores, 2000 through 2018**

		Average 3-year trend in percentiles of reading performance across PISA assessments[1]									
		10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
OECD	Australia	**-6.2**	(1.4)	**-5.5**	(1.4)	**-4.6**	(1.4)	**-3.6**	(1.4)	-2.4	(1.3)
	Austria	-0.9	(2.1)	-1.9	(2.0)	-1.9	(2.0)	-1.4	(2.0)	-1.3	(2.0)
	Belgium	1.1	(1.7)	-2.0	(1.5)	**-3.6**	(1.4)	**-2.9**	(1.3)	-1.8	(1.3)
	Canada	**-2.8**	(1.3)	**-2.7**	(1.3)	-2.0	(1.3)	-1.1	(1.3)	-0.3	(1.3)
	Chile	**8.1**	(1.6)	**7.3**	(1.5)	**7.0**	(1.5)	**6.8**	(1.5)	**6.2**	(1.5)
	Colombia	**12.0**	(2.1)	**7.4**	(2.1)	**4.3**	(2.0)	**3.8**	(1.8)	**4.5**	(1.9)
	Czech Republic	0.2	(1.5)	-1.0	(1.4)	-0.6	(1.4)	0.2	(1.4)	0.7	(1.4)
	Denmark	2.2	(1.4)	1.0	(1.3)	0.6	(1.3)	0.5	(1.4)	0.7	(1.3)
	Estonia	**3.7**	(1.9)	**4.1**	(1.6)	**5.6**	(1.5)	**7.7**	(1.5)	**10.0**	(1.6)
	Finland	**-8.6**	(1.5)	**-6.9**	(1.3)	**-4.5**	(1.3)	**-2.8**	(1.3)	-1.5	(1.3)
	France	**-4.0**	(1.5)	-2.7	(1.4)	-0.5	(1.3)	1.6	(1.3)	**3.0**	(1.3)
	Germany	**5.8**	(1.7)	**3.2**	(1.5)	2.0	(1.4)	2.1	(1.4)	1.9	(1.4)
	Greece	-0.8	(1.8)	-1.9	(1.7)	-2.0	(1.5)	-1.8	(1.4)	-1.5	(1.5)
	Hungary	-2.4	(1.5)	-2.4	(1.5)	-1.4	(1.4)	-0.3	(1.4)	0.3	(1.4)
	Iceland	**-6.5**	(1.4)	**-5.9**	(1.4)	**-5.0**	(1.3)	**-3.2**	(1.3)	-1.7	(1.4)
	Ireland	0.6	(1.5)	-0.6	(1.4)	-1.0	(1.3)	-0.7	(1.4)	-0.2	(1.5)
	Israel	2.6	(2.5)	4.3	(2.3)	**6.1**	(1.9)	**7.6**	(1.8)	**8.7**	(1.7)
	Italy	0.1	(1.6)	0.0	(1.4)	-0.1	(1.4)	0.0	(1.3)	0.4	(1.3)
	Japan	0.9	(1.8)	-0.1	(1.6)	0.2	(1.4)	0.8	(1.4)	1.8	(1.5)
	Korea	**-9.5**	(1.6)	**-6.3**	(1.4)	**-2.7**	(1.4)	0.2	(1.4)	2.6	(1.4)
	Latvia	**4.7**	(1.6)	**3.2**	(1.5)	2.2	(1.4)	1.0	(1.4)	0.3	(1.4)
	Lithuania	1.0	(1.6)	0.7	(1.6)	1.4	(1.6)	2.3	(1.6)	2.7	(1.6)
	Luxembourg	**-2.9**	(1.4)	**-3.8**	(1.3)	-2.0	(1.3)	0.7	(1.3)	**3.5**	(1.4)
	Mexico	**4.9**	(1.4)	**3.5**	(1.4)	1.8	(1.4)	0.3	(1.4)	-0.4	(1.5)
	Netherlands*	**-9.0**	(1.8)	**-6.9**	(1.6)	**-4.6**	(1.5)	-2.1	(1.5)	0.6	(1.4)
	New Zealand	**-3.2**	(1.5)	**-4.3**	(1.4)	**-4.4**	(1.3)	**-3.6**	(1.3)	**-3.3**	(1.4)
	Norway	1.4	(1.5)	0.9	(1.4)	0.5	(1.3)	1.1	(1.3)	1.3	(1.4)
	Poland	**6.4**	(1.5)	**4.8**	(1.4)	**4.1**	(1.4)	**3.8**	(1.5)	**3.3**	(1.5)
	Portugal*	**5.2**	(1.6)	**4.4**	(1.5)	**4.1**	(1.4)	**4.0**	(1.4)	**4.1**	(1.4)
	Slovak Republic	**-5.4**	(1.7)	**-4.8**	(1.6)	**-3.6**	(1.4)	-2.3	(1.5)	-0.8	(1.5)
	Slovenia	1.3	(1.5)	1.3	(1.4)	1.6	(1.3)	2.5	(1.4)	**4.3**	(1.6)
	Spain	m	m	m	m	m	m	m	m	m	m
	Sweden	**-6.4**	(1.6)	**-4.8**	(1.4)	**-3.0**	(1.4)	-1.0	(1.3)	0.1	(1.4)
	Switzerland	-1.7	(1.6)	-2.1	(1.5)	-1.8	(1.4)	-0.9	(1.5)	-0.5	(1.5)
	Turkey	**3.4**	(1.8)	2.7	(1.6)	2.5	(1.6)	2.2	(1.7)	0.7	(2.2)
	United Kingdom	2.9	(1.8)	1.1	(1.5)	1.3	(1.5)	2.1	(1.5)	2.9	(1.6)
	United States*	0.2	(2.0)	-0.6	(1.7)	-0.2	(1.6)	0.4	(1.6)	0.4	(1.5)
	OECD average-23	-0.8	(1.2)	-1.2	(1.2)	-0.9	(1.2)	-0.3	(1.2)	0.3	(1.2)
	OECD average-27	-0.3	(1.3)	-0.7	(1.3)	-0.4	(1.2)	0.3	(1.2)	0.8	(1.2)
	OECD average-29b	-1.1	(1.3)	-1.5	(1.2)	-1.1	(1.2)	-0.3	(1.2)	0.3	(1.2)
	OECD average-35a	0.0	(1.2)	-0.4	(1.2)	-0.1	(1.2)	0.7	(1.2)	1.5	(1.2)
	OECD average-35b	0.0	(1.2)	-0.5	(1.2)	-0.1	(1.2)	0.7	(1.2)	1.4	(1.2)
	OECD average-36a	0.0	(1.2)	-0.5	(1.2)	-0.1	(1.2)	0.7	(1.2)	1.4	(1.2)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.13 [2/2] **Distribution of reading scores, 2000 through 2018**

		Average 3-year trend in percentiles of reading performance across PISA assessments[1]									
		10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Partners	Albania	**14.4**	(1.7)	**12.3**	(1.6)	**9.7**	(1.5)	**7.9**	(1.4)	**7.9**	(1.4)
	Argentina	4.4	(2.4)	0.8	(2.4)	-2.5	(2.2)	-3.8	(2.0)	**-4.3**	(2.1)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m
	Belarus	m	m	m	m	m	m	m	m	m	m
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m
	Brazil	2.6	(1.5)	1.5	(1.4)	1.6	(1.4)	**3.3**	(1.4)	**4.0**	(1.4)
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m
	Bulgaria	0.9	(1.8)	-0.4	(1.8)	0.1	(1.8)	1.2	(1.8)	1.8	(1.9)
	Costa Rica	**-7.6**	(2.5)	**-8.3**	(2.2)	**-8.3**	(2.2)	**-6.1**	(2.5)	-3.7	(3.0)
	Croatia	1.4	(2.0)	0.7	(1.7)	0.4	(1.6)	1.7	(1.6)	2.9	(1.6)
	Cyprus	-0.7	(3.1)	**-12.6**	(2.4)	**-17.4**	(2.4)	**-16.8**	(2.5)	**-14.7**	(2.7)
	Dominican Republic	-9.4	(6.0)	**-15.6**	(5.9)	**-19.7**	(6.1)	**-20.8**	(6.9)	**-17.5**	(8.5)
	Georgia	**11.6**	(2.4)	**5.2**	(2.2)	0.3	(2.0)	-0.1	(2.0)	0.5	(2.4)
	Hong Kong (China)*	-1.5	(1.7)	-0.3	(1.5)	1.3	(1.4)	**3.1**	(1.3)	**4.8**	(1.3)
	Indonesia	1.2	(1.5)	0.7	(1.4)	0.7	(1.4)	0.9	(1.5)	2.1	(1.6)
	Jordan	**4.9**	(2.2)	**4.7**	(1.7)	**4.0**	(1.6)	**3.5**	(1.6)	2.6	(1.7)
	Kazakhstan	**5.0**	(1.7)	0.9	(1.6)	-2.8	(1.7)	**-6.1**	(1.8)	**-6.2**	(2.2)
	Kosovo	**22.0**	(5.2)	**10.0**	(5.0)	2.1	(4.7)	-4.3	(4.9)	-4.6	(5.2)
	Lebanon	8.4	(8.4)	3.5	(7.8)	8.4	(8.8)	8.6	(9.0)	4.0	(9.4)
	Macao (China)	-0.1	(1.5)	**3.0**	(1.4)	**6.3**	(1.3)	**9.3**	(1.5)	**11.2**	(1.4)
	Malaysia	1.1	(2.4)	-0.3	(2.2)	0.8	(2.2)	3.1	(2.2)	**5.6**	(2.4)
	Malta	5.4	(2.9)	1.1	(2.0)	0.3	(2.0)	-0.3	(2.0)	2.2	(2.0)
	Moldova	**11.1**	(2.3)	**11.1**	(2.3)	**12.8**	(2.1)	**15.9**	(2.2)	**17.0**	(2.2)
	Montenegro	**8.2**	(1.5)	**7.5**	(1.4)	**7.0**	(1.3)	**7.4**	(1.3)	**8.0**	(1.4)
	Morocco	m	m	m	m	m	m	m	m	m	m
	North Macedonia	0.5	(4.0)	1.1	(4.0)	1.1	(4.0)	0.8	(4.0)	1.6	(4.0)
	Panama	6.3	(3.7)	3.5	(2.9)	2.0	(2.8)	-0.0	(3.2)	-2.8	(3.8)
	Peru	**14.6**	(1.6)	**13.9**	(1.6)	**13.5**	(1.6)	**12.9**	(1.7)	**12.5**	(1.7)
	Philippines	m	m	m	m	m	m	m	m	m	m
	Qatar	**19.3**	(1.3)	**21.3**	(1.3)	**24.2**	(1.3)	**23.7**	(1.3)	**20.9**	(1.6)
	Romania	**5.4**	(2.3)	**6.1**	(2.4)	**7.0**	(2.2)	**8.1**	(2.1)	**10.1**	(2.2)
	Russia	**7.7**	(1.6)	**7.0**	(1.5)	**6.1**	(1.4)	**6.4**	(1.4)	**6.7**	(1.4)
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m
	Serbia	**4.8**	(1.9)	**4.9**	(1.9)	**7.0**	(1.8)	**9.0**	(1.7)	**11.3**	(1.7)
	Singapore	0.3	(1.8)	**5.0**	(1.6)	**7.4**	(1.6)	**8.6**	(1.6)	**9.5**	(1.7)
	Chinese Taipei	-3.7	(1.9)	-1.5	(1.8)	1.2	(1.6)	**4.6**	(1.6)	**7.4**	(1.8)
	Thailand	**-4.1**	(1.4)	**-4.8**	(1.4)	**-5.0**	(1.4)	**-4.0**	(1.4)	-2.6	(1.5)
	Ukraine	m	m	m	m	m	m	m	m	m	m
	United Arab Emirates	**-8.1**	(2.5)	**-6.6**	(2.4)	-2.5	(2.3)	3.8	(2.4)	**8.9**	(2.4)
	Uruguay	**8.4**	(1.6)	**3.4**	(1.5)	-0.5	(1.5)	**-3.2**	(1.5)	**-5.8**	(1.6)
	Viet Nam**	-17.1	(9.7)	**-24.2**	(7.2)	**-25.0**	(5.9)	**-22.7**	(6.2)	**-18.9**	(7.8)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Albania, Argentina, Bulgaria, Chile, Indonesia, the Republic of North Macedonia, Peru and Thailand conducted the PISA 2000 assessment in 2001, Hong Kong (China), Israel and Romania conducted the assessment in 2002, as part of PISA 2000+. Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.14 [1/2] **Distribution of mathematics scores, 2003 through 2018**

		Average 3-year trend in percentiles of mathematics performance across PISA assessments[1]									
		10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
OECD	Australia	**-7.1**	(1.1)	**-7.5**	(1.1)	**-7.5**	(1.1)	**-7.4**	(1.1)	**-6.9**	(1.2)
	Austria	-1.7	(1.5)	-1.5	(1.4)	-1.3	(1.3)	-1.8	(1.3)	-2.3	(1.3)
	Belgium	-1.1	(1.5)	**-3.4**	(1.2)	**-4.8**	(1.2)	**-6.2**	(1.1)	**-6.8**	(1.2)
	Canada	**-5.5**	(1.2)	**-5.0**	(1.2)	**-4.2**	(1.1)	**-3.5**	(1.1)	**-2.9**	(1.2)
	Chile	0.9	(1.5)	1.5	(1.4)	2.1	(1.5)	1.9	(1.6)	0.9	(1.8)
	Colombia	**7.2**	(1.8)	**5.2**	(1.6)	**4.3**	(1.5)	**4.4**	(1.6)	**4.9**	(1.5)
	Czech Republic	-2.0	(1.5)	**-2.7**	(1.3)	**-3.2**	(1.3)	**-4.8**	(1.3)	**-5.9**	(1.3)
	Denmark	1.0	(1.3)	0.2	(1.2)	-0.9	(1.1)	-2.1	(1.2)	**-3.4**	(1.2)
	Estonia	2.2	(1.4)	1.8	(1.2)	**2.4**	(1.2)	**2.6**	(1.3)	**2.6**	(1.2)
	Finland	**-9.7**	(1.2)	**-9.2**	(1.2)	**-8.5**	(1.1)	**-9.2**	(1.1)	**-9.3**	(1.1)
	France	**-3.1**	(1.4)	**-2.7**	(1.3)	-2.1	(1.2)	-2.2	(1.2)	**-2.8**	(1.2)
	Germany	2.8	(1.5)	1.1	(1.4)	-0.4	(1.3)	-1.7	(1.3)	**-2.8**	(1.2)
	Greece	0.5	(1.6)	0.0	(1.4)	0.2	(1.3)	0.1	(1.3)	-0.8	(1.4)
	Hungary	**-3.7**	(1.4)	**-3.3**	(1.2)	-2.3	(1.2)	-2.4	(1.3)	**-3.0**	(1.4)
	Iceland	**-5.6**	(1.3)	**-5.2**	(1.2)	**-4.6**	(1.1)	**-4.5**	(1.1)	**-4.1**	(1.2)
	Ireland	1.3	(1.3)	1.1	(1.3)	0.2	(1.2)	-0.8	(1.1)	-1.8	(1.2)
	Israel	**4.4**	(2.1)	**6.0**	(1.8)	**7.6**	(1.7)	**7.2**	(1.6)	**5.8**	(1.6)
	Italy	**5.2**	(1.4)	**5.6**	(1.3)	**5.9**	(1.2)	**5.4**	(1.2)	**4.6**	(1.3)
	Japan	2.9	(1.6)	1.3	(1.4)	-0.2	(1.3)	-1.5	(1.3)	-2.7	(1.5)
	Korea	**-7.3**	(1.5)	**-5.1**	(1.3)	**-3.6**	(1.3)	-2.6	(1.4)	-1.9	(1.6)
	Latvia	**3.5**	(1.5)	**2.5**	(1.3)	1.5	(1.2)	0.8	(1.2)	0.1	(1.3)
	Lithuania	-0.9	(1.4)	-1.2	(1.3)	-0.6	(1.2)	-0.6	(1.3)	-0.8	(1.7)
	Luxembourg	**-3.1**	(1.2)	**-3.2**	(1.1)	-1.9	(1.1)	-0.5	(1.1)	-0.3	(1.1)
	Mexico	**6.0**	(1.3)	**4.6**	(1.2)	**3.1**	(1.2)	1.9	(1.3)	0.7	(1.3)
	Netherlands*	**-5.2**	(1.5)	**-4.1**	(1.5)	**-3.6**	(1.3)	**-4.4**	(1.2)	**-4.1**	(1.2)
	New Zealand	**-6.0**	(1.2)	**-6.7**	(1.2)	**-7.2**	(1.1)	**-7.5**	(1.1)	**-7.9**	(1.3)
	Norway	1.8	(1.3)	2.0	(1.2)	2.1	(1.2)	1.2	(1.2)	0.5	(1.1)
	Poland	**4.4**	(1.2)	**5.1**	(1.2)	**5.3**	(1.2)	**5.0**	(1.2)	**4.7**	(1.3)
	Portugal*	2.6	(1.4)	**4.6**	(1.4)	**6.7**	(1.3)	**7.9**	(1.2)	**7.8**	(1.2)
	Slovak Republic	**-6.1**	(1.5)	**-4.6**	(1.4)	**-2.7**	(1.2)	**-2.6**	(1.2)	**-2.8**	(1.3)
	Slovenia	2.0	(1.1)	**2.8**	(1.1)	**2.8**	(1.1)	1.3	(1.0)	-0.8	(1.2)
	Spain	0.4	(1.2)	-0.1	(1.2)	0.1	(1.1)	0.0	(1.1)	-0.6	(1.1)
	Sweden	-2.0	(1.4)	-2.0	(1.2)	-1.7	(1.3)	**-2.4**	(1.2)	**-2.9**	(1.3)
	Switzerland	-1.1	(1.3)	**-2.8**	(1.3)	**-2.6**	(1.3)	**-3.0**	(1.4)	**-3.4**	(1.5)
	Turkey	**6.3**	(1.4)	**6.3**	(1.4)	**5.3**	(1.5)	**3.9**	(1.8)	-0.2	(2.6)
	United Kingdom	-0.8	(1.4)	0.6	(1.2)	2.2	(1.2)	**2.5**	(1.2)	1.9	(1.4)
	United States*	-0.1	(1.4)	-1.0	(1.3)	-1.1	(1.3)	-1.6	(1.3)	-2.3	(1.3)
	OECD average-29a	-1.0	(1.0)	-1.2	(1.0)	-1.1	(1.0)	-1.5	(1.0)	**-2.1**	(1.0)
	OECD average-30	-1.1	(1.0)	-1.2	(1.0)	-1.1	(1.0)	-1.6	(1.0)	**-2.1**	(1.0)
	OECD average-36b	-0.4	(0.9)	-0.5	(0.9)	-0.3	(0.9)	-0.7	(0.9)	-1.3	(0.9)
	OECD average-37	-0.5	(0.9)	-0.5	(0.9)	-0.4	(0.9)	-0.7	(0.9)	-1.3	(0.9)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.14 [2/2] **Distribution of mathematics scores, 2003 through 2018**

	Average 3-year trend in percentiles of mathematics performance across PISA assessments[1]									
	10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Partners										
Albania	**24.0**	(2.4)	**20.6**	(2.3)	**18.5**	(2.0)	**18.1**	(2.1)	**16.7**	(2.5)
Argentina	**5.3**	(2.4)	1.0	(2.0)	-2.2	(1.8)	**-4.1**	(1.8)	**-5.6**	(2.2)
Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m
Belarus	m	m	m	m	m	m	m	m	m	m
Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m
Brazil	**7.4**	(1.4)	**5.8**	(1.2)	**4.3**	(1.3)	**3.7**	(1.4)	2.6	(1.9)
Brunei Darussalam	m	m	m	m	m	m	m	m	m	m
B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m
Bulgaria	**6.2**	(2.1)	**5.8**	(1.9)	**5.9**	(1.9)	**5.7**	(2.0)	**5.4**	(2.5)
Costa Rica	**-5.0**	(2.2)	**-3.9**	(1.9)	-2.8	(2.2)	-1.8	(2.4)	-1.5	(2.8)
Croatia	-0.9	(1.4)	-0.9	(1.3)	-0.4	(1.3)	0.2	(1.3)	0.6	(1.5)
Cyprus	2.6	(2.4)	**4.5**	(2.2)	**7.6**	(2.2)	**7.6**	(2.2)	**5.6**	(2.3)
Dominican Republic	-6.4	(5.3)	-4.7	(4.8)	-2.3	(4.8)	-2.5	(5.3)	0.4	(7.1)
Georgia	**5.9**	(2.6)	**6.4**	(2.2)	**6.6**	(2.0)	**8.8**	(2.2)	**11.2**	(2.6)
Hong Kong (China)*	1.6	(1.8)	1.4	(1.6)	0.1	(1.3)	-0.7	(1.2)	-1.0	(1.3)
Indonesia	2.7	(1.4)	2.3	(1.2)	2.2	(1.3)	1.5	(1.5)	1.5	(1.8)
Jordan	1.6	(1.6)	1.8	(1.4)	2.3	(1.4)	2.8	(1.5)	**3.6**	(1.7)
Kazakhstan	1.3	(1.7)	**4.0**	(1.7)	**5.7**	(1.7)	**6.5**	(2.0)	**6.7**	(2.4)
Kosovo	3.1	(4.6)	3.6	(3.9)	3.6	(3.5)	3.0	(4.2)	5.3	(5.8)
Lebanon	-11.9	(7.4)	-7.0	(7.3)	-0.9	(7.1)	4.3	(7.2)	1.4	(7.7)
Macao (China)	**7.4**	(1.4)	**7.6**	(1.2)	**6.5**	(1.2)	**5.5**	(1.1)	**4.5**	(1.4)
Malaysia	**8.7**	(2.0)	**11.1**	(1.9)	**12.8**	(1.9)	**15.0**	(2.3)	**16.8**	(2.8)
Malta	3.1	(2.3)	3.7	(2.3)	**4.6**	(2.0)	**3.8**	(1.9)	2.4	(1.9)
Moldova	**5.0**	(2.4)	**6.0**	(2.3)	**8.9**	(2.2)	**11.9**	(2.4)	**13.6**	(2.9)
Montenegro	**7.8**	(1.3)	**7.0**	(1.1)	**7.5**	(1.1)	**8.1**	(1.0)	**7.8**	(1.1)
Morocco	m	m	m	m	m	m	m	m	m	m
North Macedonia	**23.7**	(4.8)	**23.7**	(4.0)	**25.4**	(3.7)	**24.1**	(4.0)	**20.2**	(4.6)
Panama	-2.1	(2.9)	-1.9	(2.4)	-1.5	(2.4)	-1.7	(2.8)	-4.0	(3.6)
Peru	**14.5**	(2.0)	**13.1**	(2.0)	**12.1**	(2.0)	**11.3**	(2.3)	**10.8**	(2.7)
Philippines	m	m	m	m	m	m	m	m	m	m
Qatar	**18.1**	(1.2)	**20.6**	(1.0)	**24.6**	(1.0)	**26.9**	(1.0)	**23.9**	(1.1)
Romania	1.2	(2.3)	2.4	(1.8)	**4.3**	(1.8)	**6.9**	(1.9)	**8.8**	(2.3)
Russia	**5.8**	(1.5)	**5.5**	(1.4)	**5.4**	(1.3)	**4.5**	(1.4)	2.8	(1.4)
Saudi Arabia	m	m	m	m	m	m	m	m	m	m
Serbia	1.2	(1.7)	1.2	(1.6)	2.2	(1.5)	**4.5**	(1.6)	**5.6**	(1.6)
Singapore	**5.9**	(1.9)	**5.3**	(1.6)	1.7	(1.5)	-2.4	(1.6)	**-5.1**	(1.8)
Chinese Taipei	-2.4	(1.8)	-2.1	(1.8)	**-3.9**	(1.5)	**-5.0**	(1.4)	**-5.2**	(1.6)
Thailand	-1.1	(1.3)	-0.4	(1.2)	0.4	(1.2)	0.8	(1.3)	1.4	(1.6)
Ukraine	m	m	m	m	m	m	m	m	m	m
United Arab Emirates	-3.7	(2.3)	0.2	(2.2)	**4.3**	(2.1)	**8.4**	(2.1)	**10.0**	(2.2)
Uruguay	**3.1**	(1.3)	0.0	(1.3)	**-2.9**	(1.3)	**-4.3**	(1.3)	**-5.4**	(1.4)
Viet Nam**	-12.9	(9.1)	**-18.2**	(7.3)	**-18.3**	(6.8)	**-17.7**	(7.3)	**-21.2**	(9.8)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted.

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.15 [1/2] **Distribution of science scores, 2006 through 2018**

		Average 3-year trend in percentiles of science performance across PISA assessments[1]									
		10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
OECD	Australia	**-7.6**	(1.4)	**-7.7**	(1.3)	**-6.6**	(1.3)	**-5.9**	(1.3)	**-6.0**	(1.4)
	Austria	**-4.7**	(2.0)	**-6.4**	(1.9)	**-6.3**	(1.7)	**-5.7**	(1.6)	**-5.1**	(1.6)
	Belgium	-2.2	(1.9)	**-3.5**	(1.5)	**-3.3**	(1.4)	**-3.3**	(1.3)	**-2.5**	(1.3)
	Canada	**-4.3**	(1.4)	**-4.3**	(1.3)	**-3.8**	(1.3)	**-3.0**	(1.3)	-2.0	(1.4)
	Chile	1.9	(1.6)	1.7	(1.5)	1.6	(1.6)	0.8	(1.9)	-0.7	(1.8)
	Colombia	**8.0**	(1.7)	**5.5**	(1.7)	**5.0**	(1.5)	**6.1**	(1.7)	**7.3**	(1.7)
	Czech Republic	-3.2	(1.8)	**-3.9**	(1.6)	**-4.4**	(1.5)	**-4.5**	(1.5)	**-4.9**	(1.6)
	Denmark	0.2	(1.7)	-0.1	(1.5)	-0.3	(1.4)	-0.8	(1.4)	-1.2	(1.5)
	Estonia	-1.3	(1.7)	-0.9	(1.4)	0.4	(1.3)	1.5	(1.4)	2.1	(1.5)
	Finland	**-15.5**	(1.6)	**-12.8**	(1.4)	**-10.3**	(1.3)	**-8.1**	(1.3)	**-7.2**	(1.4)
	France	0.7	(1.8)	-1.0	(1.7)	-1.4	(1.6)	-1.4	(1.5)	-1.7	(1.5)
	Germany	**-4.2**	(2.1)	**-4.4**	(1.8)	**-3.8**	(1.6)	**-3.3**	(1.5)	-2.6	(1.4)
	Greece	**-5.3**	(2.0)	**-6.6**	(1.8)	**-6.4**	(1.6)	**-6.1**	(1.5)	**-6.4**	(1.5)
	Hungary	**-10.6**	(1.8)	**-10.1**	(1.5)	**-7.3**	(1.5)	**-5.0**	(1.4)	**-3.6**	(1.5)
	Iceland	**-3.8**	(1.5)	**-5.7**	(1.4)	**-6.1**	(1.3)	**-6.4**	(1.3)	**-6.1**	(1.4)
	Ireland	-0.7	(1.6)	-2.1	(1.5)	**-3.6**	(1.4)	**-4.5**	(1.5)	**-5.0**	(1.6)
	Israel	2.0	(1.9)	2.1	(1.9)	3.3	(1.8)	**3.5**	(1.7)	2.9	(1.7)
	Italy	-0.9	(1.5)	-1.3	(1.4)	-2.4	(1.4)	**-3.3**	(1.3)	**-4.3**	(1.5)
	Japan	2.3	(2.0)	-0.2	(1.7)	-1.5	(1.5)	-2.0	(1.4)	-2.2	(1.4)
	Korea	**-7.6**	(1.8)	**-5.3**	(1.6)	-2.5	(1.5)	-0.7	(1.5)	1.0	(1.7)
	Latvia	-1.4	(1.7)	-1.4	(1.5)	-1.1	(1.4)	-0.3	(1.4)	0.1	(1.4)
	Lithuania	**-3.7**	(1.5)	**-3.6**	(1.4)	**-3.3**	(1.4)	-2.1	(1.4)	-1.5	(1.5)
	Luxembourg	-1.4	(1.5)	**-3.7**	(1.3)	**-3.2**	(1.2)	-1.4	(1.2)	-0.6	(1.3)
	Mexico	**4.5**	(1.6)	2.7	(1.4)	1.5	(1.4)	0.4	(1.4)	-0.2	(1.5)
	Netherlands*	**-8.5**	(2.0)	**-7.6**	(1.8)	**-5.8**	(1.7)	**-3.8**	(1.4)	**-2.9**	(1.5)
	New Zealand	**-5.1**	(1.6)	**-6.0**	(1.5)	**-6.5**	(1.5)	**-7.1**	(1.4)	**-7.1**	(1.4)
	Norway	-2.7	(1.7)	-0.5	(1.5)	1.3	(1.4)	1.6	(1.4)	1.9	(1.5)
	Poland	1.0	(1.4)	1.5	(1.4)	2.2	(1.4)	2.6	(1.4)	3.0	(1.5)
	Portugal*	1.7	(1.8)	3.2	(1.6)	**4.3**	(1.5)	**5.7**	(1.4)	**6.0**	(1.5)
	Slovak Republic	**-10.0**	(1.7)	**-9.4**	(1.4)	**-7.7**	(1.4)	**-7.0**	(1.5)	**-6.2**	(1.5)
	Slovenia	-0.2	(1.5)	-0.5	(1.3)	-1.8	(1.2)	**-3.9**	(1.2)	**-5.0**	(1.4)
	Spain	-0.9	(1.4)	-1.0	(1.3)	-0.8	(1.3)	-0.3	(1.3)	-0.2	(1.3)
	Sweden	**-3.8**	(1.9)	-2.3	(1.6)	-0.6	(1.5)	0.4	(1.4)	0.5	(1.5)
	Switzerland	**-3.9**	(1.7)	**-5.7**	(1.6)	**-5.1**	(1.5)	**-4.3**	(1.6)	-3.3	(1.8)
	Turkey	**4.8**	(1.5)	**5.7**	(1.4)	**7.1**	(1.5)	**7.4**	(1.8)	**5.1**	(2.4)
	United Kingdom	-1.0	(1.6)	-1.8	(1.5)	-2.6	(1.3)	**-3.3**	(1.4)	**-4.2**	(1.4)
	United States*	3.6	(1.9)	3.4	(1.8)	2.5	(1.7)	1.0	(1.8)	-0.2	(1.7)
	OECD average-36b	**-2.2**	(1.1)	**-2.4**	(1.1)	-2.0	(1.1)	-1.7	(1.1)	-1.6	(1.1)
	OECD average-37	**-2.3**	(1.1)	**-2.5**	(1.1)	**-2.1**	(1.1)	-1.8	(1.1)	-1.7	(1.1)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Table I.B1.15 [2/2] **Distribution of science scores, 2006 through 2018**

		Average 3-year trend in percentiles of science performance across PISA assessments[1]									
		10th percentile		25th percentile		Median (50th percentile)		75th percentile		90th percentile	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Partners	**Albania**	**19.7**	(2.4)	**13.4**	(2.2)	**9.2**	(2.2)	**5.6**	(2.2)	4.1	(2.5)
	Argentina	**8.3**	(2.3)	**4.0**	(1.9)	1.5	(1.8)	0.5	(1.8)	0.0	(1.9)
	Baku (Azerbaijan)	m	m	m	m	m	m	m	m	m	m
	Belarus	m	m	m	m	m	m	m	m	m	m
	Bosnia and Herzegovina	m	m	m	m	m	m	m	m	m	m
	Brazil	1.2	(1.4)	0.9	(1.2)	1.8	(1.3)	**3.4**	(1.5)	**4.0**	(1.8)
	Brunei Darussalam	m	m	m	m	m	m	m	m	m	m
	B-S-J-Z (China)	m	m	m	m	m	m	m	m	m	m
	Bulgaria	2.0	(2.2)	-0.3	(2.0)	-2.0	(2.2)	-3.3	(2.3)	**-4.6**	(2.3)
	Costa Rica	**-5.2**	(2.3)	**-6.4**	(2.4)	**-6.7**	(2.3)	**-5.8**	(2.4)	-5.5	(3.1)
	Croatia	**-7.4**	(1.6)	**-6.8**	(1.5)	**-5.8**	(1.5)	**-4.3**	(1.4)	-2.9	(1.5)
	Cyprus	3.1	(2.8)	-0.6	(2.5)	-1.0	(2.5)	0.6	(2.6)	0.7	(2.6)
	Dominican Republic	6.2	(4.2)	4.9	(3.8)	3.5	(4.3)	3.2	(5.1)	2.0	(7.0)
	Georgia	**10.6**	(2.2)	**6.7**	(2.2)	4.0	(2.2)	3.3	(2.2)	3.4	(2.5)
	Hong Kong (China)*	**-5.4**	(1.9)	**-6.3**	(1.5)	**-8.1**	(1.4)	**-9.4**	(1.3)	**-9.6**	(1.4)
	Indonesia	**3.0**	(1.5)	2.7	(1.6)	2.3	(1.7)	2.2	(2.1)	1.9	(2.8)
	Jordan	1.1	(1.7)	1.0	(1.5)	0.8	(1.4)	0.3	(1.5)	-0.1	(1.6)
	Kazakhstan	2.2	(1.9)	-1.0	(1.8)	**-4.3**	(1.7)	**-7.1**	(1.9)	**-6.6**	(2.5)
	Kosovo	-3.5	(3.6)	**-8.2**	(3.1)	**-14.4**	(2.9)	**-20.3**	(3.1)	**-24.1**	(4.8)
	Lebanon	-10.6	(5.5)	-6.9	(5.4)	-2.0	(6.2)	2.5	(7.2)	2.6	(7.1)
	Macao (China)	**6.0**	(1.3)	**7.6**	(1.2)	**8.4**	(1.2)	**9.1**	(1.2)	**9.7**	(1.3)
	Malaysia	**6.5**	(2.2)	**6.0**	(1.9)	**6.1**	(2.0)	**6.8**	(2.1)	**7.5**	(2.5)
	Malta	2.6	(2.2)	-0.7	(2.1)	-2.2	(2.0)	-3.8	(2.2)	**-4.3**	(2.1)
	Moldova	**5.9**	(2.4)	4.1	(2.4)	4.4	(2.3)	**7.1**	(2.1)	**8.6**	(2.4)
	Montenegro	1.0	(1.3)	1.4	(1.2)	1.3	(1.2)	1.7	(1.2)	2.6	(1.4)
	Morocco	m	m	m	m	m	m	m	m	m	m
	North Macedonia	**18.9**	(3.8)	**23.9**	(3.1)	**29.1**	(3.3)	**36.1**	(4.3)	**36.6**	(4.8)
	Panama	-0.4	(3.2)	-3.3	(3.0)	-4.1	(2.8)	-5.4	(2.9)	-5.5	(3.5)
	Peru	**17.3**	(2.2)	**13.5**	(2.0)	**11.7**	(2.0)	**11.0**	(2.3)	**10.3**	(2.7)
	Philippines	m	m	m	m	m	m	m	m	m	m
	Qatar	**11.3**	(1.1)	**14.3**	(1.1)	**19.2**	(1.1)	**23.1**	(1.2)	**22.2**	(1.4)
	Romania	0.1	(1.8)	0.8	(1.8)	2.1	(1.9)	3.3	(1.9)	**4.5**	(2.1)
	Russia	2.5	(1.8)	1.4	(1.6)	0.6	(1.6)	-0.6	(1.5)	-1.9	(1.6)
	Saudi Arabia	m	m	m	m	m	m	m	m	m	m
	Serbia	-1.8	(1.6)	-1.3	(1.5)	0.0	(1.5)	1.9	(1.6)	**4.4**	(1.7)
	Singapore	**4.4**	(2.1)	**5.6**	(1.7)	**4.4**	(1.8)	1.8	(1.7)	-0.9	(1.8)
	Chinese Taipei	**-4.6**	(1.7)	-3.2	(1.8)	-2.5	(1.5)	-0.9	(1.5)	0.5	(1.4)
	Thailand	-0.5	(1.5)	-0.9	(1.4)	0.1	(1.4)	1.6	(1.6)	2.3	(1.7)
	Ukraine	m	m	m	m	m	m	m	m	m	m
	United Arab Emirates	**-6.5**	(2.4)	**-6.2**	(2.4)	-3.7	(2.4)	0.9	(2.4)	3.0	(2.4)
	Uruguay	**4.0**	(1.7)	1.1	(1.5)	-0.4	(1.4)	-1.0	(1.5)	-1.9	(1.6)
	Viet Nam**	-0.7	(8.4)	-8.5	(7.0)	-7.9	(6.2)	-4.2	(6.3)	-0.9	(9.1)

1. The average 3-year trend is the average change, per 3-year-period, between the earliest available measurement in PISA and PISA 2018, calculated by a linear regression.

*PISA 2018 data did not meet the PISA technical standards but were accepted as largely comparable (see Annexes A2 and A4).

**The data for Viet Nam have not yet been fully validated. Due to a lack of consistency in the response pattern of some performance data, the OECD cannot yet assure full international comparability of the results (see Annexes A2 and A4).

Notes: Values that are statistically significant are indicated in bold (see Annex A3).

Costa Rica, Georgia, Malta and Moldova conducted the PISA 2009 assessment in 2010 as part of PISA 2009+.

Estimates of the average 3-year trend and the curvilinear trend for these countries consider the year in which the assessment was conducted

The full version of this table is available on line, at the *StatLink* below.

StatLink https://doi.org/10.1787/888934029090

Annex B1 **List of tables available on line**

ANNEX B2
Results for regions within countries

Table I.B2.9 [1/2] **Mean score and variation in reading performance, by region**

OECD	Mean score		Standard deviation		Percentiles													
					5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Belgium																		
*Flemish Community**	502	(3.4)	104	(1.9)	323	(6.3)	359	(6.4)	429	(5.1)	506	(4.1)	579	(3.1)	633	(3.2)	664	(4.1)
French Community	481	(3.0)	100	(1.8)	310	(6.9)	345	(4.7)	412	(4.2)	487	(3.8)	553	(3.6)	608	(3.9)	636	(3.9)
German-speaking Community	483	(4.6)	91	(3.5)	324	(12.0)	360	(12.1)	423	(9.0)	485	(7.3)	548	(7.2)	602	(9.3)	628	(9.8)
Canada																		
Alberta	532	(4.3)	101	(2.3)	357	(8.9)	396	(7.6)	464	(5.7)	537	(4.6)	604	(4.8)	659	(5.2)	689	(6.6)
British Columbia	519	(4.5)	104	(2.1)	342	(8.2)	380	(6.7)	448	(6.1)	524	(5.7)	595	(4.8)	649	(4.3)	680	(5.9)
Manitoba	494	(3.4)	99	(1.6)	329	(6.4)	366	(5.1)	427	(4.5)	497	(4.0)	562	(4.9)	621	(5.7)	655	(6.1)
New Brunswick	489	(3.5)	103	(2.4)	316	(7.1)	352	(5.9)	419	(5.3)	490	(4.4)	564	(5.9)	621	(7.8)	656	(9.4)
Newfoundland and Labrador	512	(4.3)	99	(2.8)	344	(9.5)	383	(7.7)	442	(6.6)	514	(5.2)	581	(6.2)	638	(7.4)	671	(9.5)
Nova Scotia	516	(3.9)	102	(2.3)	343	(8.3)	383	(6.1)	447	(5.4)	519	(4.2)	586	(4.4)	645	(7.8)	679	(7.5)
Ontario	524	(3.5)	101	(1.6)	352	(5.6)	390	(5.0)	455	(4.7)	528	(4.2)	596	(4.0)	650	(4.3)	681	(5.4)
Prince Edward Island	503	(8.3)	103	(5.6)	325	(26.6)	364	(18.4)	435	(13.2)	509	(8.1)	574	(11.0)	635	(10.9)	662	(12.9)
Québec	519	(3.5)	94	(1.8)	358	(5.8)	396	(4.8)	457	(4.2)	523	(4.0)	586	(4.3)	637	(4.4)	666	(4.5)
Saskatchewan	499	(3.0)	95	(2.2)	338	(6.9)	376	(6.2)	436	(4.3)	501	(3.8)	565	(4.0)	621	(4.7)	651	(7.0)
Colombia																		
Bogotá	455	(5.4)	90	(2.9)	310	(7.5)	339	(7.2)	392	(6.0)	451	(6.0)	518	(7.1)	575	(7.9)	606	(8.3)
Italy																		
Bolzano	495	(3.3)	89	(2.0)	344	(7.6)	377	(6.2)	435	(4.2)	500	(4.2)	560	(4.7)	607	(5.9)	633	(7.6)
Sardegna	462	(4.1)	92	(2.1)	310	(6.4)	340	(6.0)	397	(6.0)	464	(4.8)	527	(4.6)	580	(5.1)	612	(6.6)
Toscana	482	(4.0)	94	(2.4)	315	(9.9)	355	(8.3)	420	(5.6)	490	(5.3)	550	(4.3)	597	(4.2)	625	(5.9)
Trento	496	(2.3)	93	(2.0)	340	(6.1)	371	(5.0)	430	(3.9)	499	(3.7)	564	(5.3)	615	(4.9)	643	(6.7)
Spain																		
Andalusia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Aragon	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Asturias	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Balearic Islands	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Basque Country	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Canary Islands	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Cantabria	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Castile and León	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Castile-La Mancha	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Catalonia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Ceuta	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Comunidad Valenciana	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Extremadura	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Galicia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
La Rioja	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Madrid	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Melilla	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Murcia	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Navarre	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
United Kingdom																		
England	505	(3.0)	101	(1.5)	334	(5.1)	372	(5.2)	436	(3.9)	508	(3.2)	577	(3.5)	634	(4.1)	666	(4.5)
Northern Ireland	501	(4.0)	98	(2.2)	332	(7.0)	368	(5.8)	434	(5.3)	506	(5.0)	571	(5.2)	623	(5.6)	655	(6.0)
*Scotland**	504	(3.0)	95	(1.9)	349	(5.5)	383	(3.6)	439	(3.5)	503	(3.7)	571	(4.2)	627	(4.7)	657	(5.9)
Wales	483	(4.0)	97	(1.6)	322	(5.7)	359	(5.8)	417	(4.8)	484	(4.3)	552	(4.2)	608	(4.5)	640	(6.1)

* PISA adjudicated region.

Note: See Table I.B1.4 for national data.

StatLink https://doi.org/10.1787/888934029109

Table I.B2.9 [2/2] **Mean score and variation in reading performance, by region**

		Mean score		Standard deviation		Percentiles													
						5th		10th		25th		Median (50th)		75th		90th		95th	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners	**Argentina**																		
	*CABA**	454	(5.4)	93	(2.3)	294	(8.1)	328	(8.0)	390	(7.3)	459	(6.4)	522	(5.6)	572	(5.2)	598	(6.3)
	*Córdoba**	427	(4.5)	91	(2.5)	274	(5.8)	305	(5.5)	362	(5.6)	431	(5.3)	493	(5.5)	543	(6.4)	571	(6.9)
	*PBA**	413	(5.8)	97	(2.3)	251	(7.3)	286	(7.0)	346	(6.1)	414	(7.1)	482	(7.0)	539	(6.9)	571	(7.2)
	*Tucumán**	389	(5.0)	96	(2.5)	237	(6.8)	267	(5.5)	319	(5.1)	387	(6.6)	457	(7.3)	516	(6.5)	548	(7.0)
	Brazil																		
	North	392	(6.9)	91	(2.6)	254	(7.6)	280	(8.6)	328	(7.1)	383	(8.7)	451	(9.6)	514	(11.8)	558	(10.9)
	Northeast	389	(4.2)	99	(3.1)	243	(4.9)	269	(4.8)	316	(4.3)	378	(5.2)	456	(5.8)	525	(7.4)	566	(8.7)
	South	432	(6.3)	97	(3.0)	278	(8.3)	308	(7.5)	362	(6.9)	430	(7.5)	500	(9.3)	562	(8.6)	596	(8.9)
	Southeast	424	(3.0)	98	(1.6)	264	(3.8)	296	(3.3)	353	(3.3)	421	(3.6)	494	(4.2)	555	(4.9)	588	(5.9)
	Middle-West	425	(9.1)	103	(6.1)	269	(10.8)	294	(10.0)	348	(10.1)	418	(11.0)	493	(12.1)	565	(14.9)	608	(22.2)
	Indonesia																		
	DI Yogyakarta	414	(5.8)	83	(3.2)	283	(6.0)	309	(5.8)	352	(5.3)	411	(7.1)	472	(8.6)	527	(10.4)	557	(11.2)
	DKI Jakarta	412	(7.0)	83	(4.9)	288	(6.3)	310	(4.6)	352	(5.6)	404	(6.5)	468	(11.0)	528	(15.6)	562	(18.6)
	Kazakhstan																		
	Akmola region	395	(4.5)	78	(2.0)	271	(6.0)	294	(6.1)	339	(6.0)	394	(5.6)	449	(6.6)	498	(6.2)	528	(7.7)
	Aktobe region	381	(4.3)	69	(2.2)	278	(6.9)	298	(5.4)	332	(4.7)	376	(5.1)	422	(5.4)	470	(6.0)	503	(9.7)
	Almaty	424	(7.8)	86	(4.5)	296	(6.8)	320	(7.0)	363	(7.8)	416	(8.1)	479	(10.3)	542	(16.1)	579	(16.1)
	Almaty region	360	(4.4)	67	(2.2)	254	(5.6)	273	(5.6)	312	(5.4)	359	(5.4)	402	(5.7)	449	(7.7)	479	(9.3)
	Astana	428	(7.4)	81	(3.3)	302	(8.2)	327	(8.7)	371	(7.4)	422	(8.6)	480	(10.1)	538	(10.3)	570	(12.4)
	Atyrau region	344	(4.4)	68	(2.8)	241	(5.3)	261	(5.2)	295	(5.6)	340	(5.4)	385	(5.1)	430	(6.2)	461	(8.5)
	East-Kazakhstan region	405	(6.4)	78	(3.7)	285	(11.2)	306	(10.3)	351	(8.4)	401	(6.0)	456	(8.3)	509	(9.3)	539	(11.4)
	Karagandy region	422	(6.8)	84	(3.2)	298	(8.7)	321	(6.2)	363	(6.2)	415	(7.8)	477	(10.2)	538	(12.1)	573	(11.1)
	Kostanay region	417	(5.1)	77	(2.2)	296	(8.8)	317	(7.8)	363	(6.6)	415	(6.3)	469	(6.1)	521	(5.6)	549	(5.5)
	Kyzyl-Orda region	366	(2.8)	59	(2.1)	277	(4.7)	295	(4.7)	326	(4.1)	363	(3.9)	402	(3.6)	442	(5.0)	468	(5.8)
	Mangistau region	361	(5.8)	70	(3.2)	259	(9.1)	280	(6.8)	313	(5.7)	353	(6.2)	399	(7.2)	455	(11.9)	492	(12.6)
	North-Kazakhstan region	413	(5.0)	78	(2.2)	289	(7.0)	313	(5.5)	359	(5.7)	411	(5.7)	464	(7.1)	512	(6.5)	545	(9.6)
	Pavlodar region	391	(6.5)	82	(2.4)	265	(9.0)	288	(7.8)	331	(7.5)	386	(7.8)	446	(7.6)	501	(8.3)	533	(9.3)
	South-Kazakhstan region	368	(3.5)	64	(2.3)	266	(6.2)	288	(5.9)	326	(5.2)	367	(4.2)	408	(4.1)	447	(5.8)	472	(6.8)
	West-Kazakhstan region	378	(4.9)	71	(2.3)	269	(6.6)	291	(5.3)	328	(5.4)	373	(5.8)	423	(6.2)	474	(8.0)	506	(9.5)
	Zhambyl region	369	(3.6)	63	(2.2)	277	(5.5)	295	(4.3)	325	(3.3)	363	(4.0)	406	(4.8)	451	(7.0)	483	(8.8)
	Russia																		
	*Moscow region**	486	(4.7)	92	(2.1)	327	(8.4)	364	(7.9)	424	(6.2)	488	(5.3)	552	(4.4)	604	(4.6)	632	(4.8)
	*Republic of Tatarstan**	463	(3.1)	91	(1.6)	313	(4.8)	345	(3.8)	399	(3.6)	462	(3.5)	526	(3.5)	581	(4.4)	613	(5.2)

* PISA adjudicated region.

Note: See Table I.B1.4 for national data.

StatLink http://dx.doi.org/10.1787/888934029109

Table I.B2.10 [1/2] **Mean score and variation in mathematics performance, by region**

OECD

	Mean score		Standard deviation		Percentiles 5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Belgium																		
*Flemish Community**	518	(3.3)	96	(2.3)	350	(6.3)	384	(6.0)	450	(5.5)	525	(3.6)	589	(3.0)	638	(3.6)	665	(3.9)
French Community	495	(2.8)	93	(2.0)	338	(5.4)	368	(4.5)	427	(4.0)	500	(3.6)	564	(3.6)	614	(4.6)	639	(5.2)
German-speaking Community	505	(5.2)	79	(3.9)	363	(12.7)	396	(11.8)	452	(7.5)	510	(7.0)	564	(7.3)	604	(9.3)	626	(13.0)
Canada																		
Alberta	511	(5.1)	91	(2.9)	356	(9.1)	392	(8.3)	450	(7.0)	513	(5.6)	575	(5.7)	626	(5.9)	655	(7.4)
British Columbia	504	(5.2)	94	(2.6)	350	(7.9)	382	(6.8)	441	(6.0)	505	(5.6)	569	(5.7)	624	(6.9)	657	(7.8)
Manitoba	482	(3.7)	88	(1.9)	337	(7.1)	368	(5.3)	421	(4.5)	483	(4.5)	542	(4.2)	594	(5.9)	624	(6.1)
New Brunswick	491	(5.7)	92	(2.6)	338	(8.3)	373	(7.2)	428	(6.4)	492	(6.4)	555	(7.2)	609	(9.2)	638	(10.8)
Newfoundland and Labrador	488	(6.5)	85	(2.4)	351	(10.4)	382	(8.7)	431	(5.9)	487	(6.5)	546	(8.4)	599	(10.6)	629	(11.4)
Nova Scotia	494	(6.3)	88	(2.2)	349	(8.3)	380	(8.3)	433	(6.7)	494	(6.3)	555	(6.7)	608	(8.9)	640	(11.2)
Ontario	513	(4.4)	91	(2.0)	361	(5.9)	394	(5.2)	450	(4.7)	513	(5.2)	577	(5.5)	629	(5.2)	660	(6.7)
Prince Edward Island	487	(11.1)	91	(4.4)	332	(23.0)	369	(16.4)	423	(11.6)	491	(11.7)	551	(14.2)	601	(15.2)	630	(18.1)
Québec	532	(3.6)	93	(2.2)	374	(6.8)	411	(6.2)	472	(4.8)	536	(4.0)	596	(4.1)	648	(4.2)	679	(5.2)
Saskatchewan	485	(5.0)	82	(2.2)	348	(6.5)	378	(5.4)	430	(5.8)	487	(5.8)	543	(5.4)	589	(5.7)	618	(6.9)
Colombia																		
Bogotá	430	(5.0)	81	(2.8)	301	(5.7)	326	(5.2)	372	(5.2)	428	(5.2)	484	(7.3)	536	(8.0)	567	(9.8)
Italy																		
Bolzano	521	(3.4)	84	(2.1)	378	(7.5)	410	(5.7)	465	(5.0)	525	(4.5)	582	(4.8)	626	(5.6)	652	(6.2)
Sardegna	467	(4.0)	84	(2.3)	331	(7.5)	358	(6.2)	409	(5.1)	468	(4.7)	526	(5.1)	576	(4.9)	602	(5.4)
Toscana	496	(4.3)	85	(2.1)	350	(6.4)	382	(6.1)	437	(5.3)	501	(5.7)	556	(5.5)	602	(5.7)	629	(5.6)
Trento	518	(2.8)	85	(2.1)	374	(9.2)	406	(5.6)	461	(4.6)	520	(4.2)	578	(3.9)	627	(5.3)	653	(5.8)
Spain																		
Andalusia	467	(4.2)	88	(2.1)	318	(6.4)	353	(6.4)	408	(4.7)	470	(5.0)	529	(5.9)	580	(7.3)	609	(6.0)
Aragon	497	(5.9)	89	(2.1)	343	(9.9)	377	(7.0)	437	(6.7)	501	(6.4)	560	(5.8)	607	(7.3)	634	(8.4)
Asturias	491	(5.0)	88	(2.1)	340	(9.2)	375	(7.2)	431	(5.3)	494	(4.8)	554	(6.0)	602	(6.0)	629	(6.5)
Balearic Islands	483	(5.2)	84	(1.9)	340	(8.7)	371	(7.2)	426	(6.2)	487	(5.6)	541	(5.3)	587	(5.6)	614	(6.1)
Basque Country	499	(3.5)	85	(1.6)	350	(5.5)	385	(5.1)	443	(4.7)	505	(3.9)	560	(3.8)	604	(3.9)	630	(4.7)
Canary Islands	460	(4.5)	85	(2.1)	321	(6.5)	349	(7.1)	400	(5.7)	461	(4.9)	520	(5.6)	571	(7.3)	600	(5.8)
Cantabria	499	(7.6)	85	(2.5)	354	(10.1)	387	(8.6)	444	(8.7)	503	(7.5)	558	(7.9)	605	(8.4)	631	(9.4)
Castile and León	502	(4.7)	88	(2.1)	349	(10.4)	385	(8.3)	445	(6.4)	508	(5.2)	566	(4.7)	612	(5.3)	637	(6.7)
Castile-La Mancha	479	(5.1)	88	(1.6)	332	(7.3)	362	(6.8)	416	(6.0)	482	(5.9)	544	(5.9)	591	(5.3)	617	(6.1)
Catalonia	490	(3.9)	88	(2.4)	340	(8.4)	374	(5.9)	430	(5.0)	492	(4.6)	552	(5.4)	601	(6.3)	630	(7.9)
Ceuta	411	(12.2)	84	(3.7)	282	(16.5)	308	(14.7)	352	(16.5)	407	(12.9)	467	(13.4)	525	(12.4)	557	(17.9)
Comunidad Valenciana	473	(4.6)	84	(1.8)	333	(6.1)	365	(6.3)	417	(4.5)	475	(5.4)	532	(5.4)	581	(5.7)	608	(7.5)
Extremadura	470	(6.6)	86	(2.1)	323	(10.0)	357	(7.4)	412	(7.1)	472	(7.5)	530	(6.6)	578	(8.3)	605	(8.3)
Galicia	498	(4.3)	87	(2.3)	344	(10.3)	381	(7.1)	442	(5.3)	504	(4.5)	560	(4.3)	606	(6.2)	631	(6.7)
La Rioja	497	(9.8)	92	(3.5)	338	(10.1)	374	(10.7)	437	(8.6)	502	(11.0)	562	(11.7)	613	(11.9)	643	(14.0)
Madrid	486	(3.2)	89	(1.5)	334	(6.2)	367	(5.0)	425	(4.1)	490	(3.5)	549	(3.6)	598	(3.8)	626	(3.9)
Melilla	432	(10.4)	85	(4.4)	299	(17.4)	323	(13.3)	371	(11.7)	428	(12.1)	492	(15.2)	548	(12.9)	572	(17.5)
Murcia	474	(5.7)	92	(2.2)	319	(8.8)	353	(7.7)	410	(7.2)	477	(6.0)	539	(7.0)	592	(7.2)	618	(7.7)
Navarre	503	(8.4)	87	(2.5)	355	(9.0)	387	(8.4)	444	(9.5)	506	(9.0)	564	(9.0)	612	(9.6)	640	(10.6)
United Kingdom																		
England	504	(3.0)	93	(1.7)	347	(5.1)	383	(4.9)	441	(3.4)	506	(3.2)	569	(3.6)	623	(3.7)	654	(4.5)
Northern Ireland	492	(4.2)	85	(2.5)	343	(7.8)	377	(6.4)	434	(5.5)	496	(4.4)	553	(5.6)	600	(5.3)	626	(6.8)
*Scotland**	489	(3.9)	95	(2.9)	331	(9.0)	367	(6.0)	425	(4.9)	490	(4.3)	556	(4.7)	610	(5.7)	642	(6.7)
Wales	487	(3.9)	82	(1.5)	350	(5.5)	381	(5.4)	431	(4.2)	488	(4.4)	545	(4.3)	592	(4.4)	619	(5.5)

* PISA adjudicated region.

Note: See Table I.B1.5 for national data.

StatLink http://dx.doi.org/10.1787/888934029109

Table I.B2.10 [2/2] **Mean score and variation in mathematics performance, by region**

	Mean score		Standard deviation		Percentiles													
					5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners																		
Argentina																		
*CABA**	434	(4.7)	80	(2.4)	300	(7.7)	331	(6.7)	380	(6.1)	435	(5.7)	490	(4.9)	537	(5.4)	565	(6.4)
*Córdoba**	400	(4.5)	80	(2.8)	269	(6.7)	296	(5.7)	345	(5.3)	400	(4.8)	456	(5.6)	504	(6.8)	532	(8.7)
*PBA**	387	(5.1)	84	(2.3)	254	(6.7)	282	(5.7)	330	(6.0)	384	(5.7)	443	(7.2)	498	(6.8)	529	(6.9)
*Tucumán**	364	(5.1)	81	(2.7)	237	(5.6)	262	(5.7)	308	(5.0)	362	(5.9)	418	(6.8)	472	(7.8)	502	(10.3)
Brazil																		
North	366	(7.1)	83	(4.5)	240	(10.1)	266	(9.1)	309	(7.0)	359	(7.3)	415	(10.0)	478	(12.3)	519	(21.4)
Northeast	363	(3.7)	86	(3.5)	236	(4.7)	261	(4.2)	303	(4.2)	356	(4.1)	417	(5.1)	478	(7.1)	515	(11.3)
South	401	(5.3)	87	(3.0)	266	(8.4)	293	(7.0)	338	(6.4)	396	(7.1)	460	(7.8)	520	(8.1)	554	(8.4)
Southeast	392	(3.1)	86	(2.2)	260	(3.6)	286	(3.4)	332	(3.9)	387	(3.3)	448	(4.2)	506	(5.7)	542	(7.0)
Middle-West	396	(8.4)	92	(7.0)	258	(10.8)	284	(9.3)	332	(8.2)	389	(8.9)	451	(10.9)	517	(16.4)	562	(26.5)
Indonesia																		
DI Yogyakarta	430	(6.5)	86	(3.6)	293	(6.7)	322	(6.0)	370	(6.6)	426	(7.4)	488	(9.0)	545	(10.6)	575	(12.7)
DKI Jakarta	421	(7.7)	87	(5.2)	290	(5.7)	315	(4.5)	359	(4.7)	415	(8.9)	478	(11.5)	536	(15.7)	574	(20.0)
Kazakhstan																		
Akmola region	411	(6.5)	87	(3.1)	271	(9.7)	301	(8.3)	352	(8.3)	411	(7.1)	470	(7.8)	522	(8.8)	556	(10.8)
Aktobe region	420	(6.2)	80	(3.2)	286	(12.4)	319	(8.4)	366	(8.0)	421	(6.6)	474	(7.1)	522	(7.3)	550	(6.8)
Almaty	448	(7.3)	90	(3.4)	302	(10.8)	334	(9.0)	386	(8.0)	447	(8.6)	508	(9.0)	567	(11.0)	603	(12.4)
Almaty region	399	(5.2)	80	(2.4)	268	(8.3)	297	(6.9)	345	(5.9)	399	(5.7)	453	(6.7)	502	(7.2)	530	(9.5)
Astana	450	(7.7)	89	(3.5)	309	(10.7)	339	(9.5)	390	(9.0)	447	(8.5)	509	(9.4)	568	(10.1)	602	(10.8)
Atyrau region	382	(7.1)	80	(3.0)	251	(11.5)	281	(9.3)	330	(7.7)	382	(7.8)	434	(7.9)	481	(9.4)	513	(9.2)
East-Kazakhstan region	437	(7.3)	88	(4.2)	293	(14.1)	327	(11.0)	379	(8.9)	438	(7.4)	495	(8.5)	549	(10.8)	580	(11.4)
Karagandy region	446	(7.4)	90	(4.3)	307	(8.2)	335	(7.3)	384	(7.1)	442	(7.9)	503	(8.6)	562	(13.3)	602	(16.5)
Kostanay region	448	(6.6)	86	(4.1)	303	(14.2)	338	(11.3)	391	(9.3)	450	(7.7)	506	(6.5)	557	(7.0)	585	(7.6)
Kyzyl-Orda region	419	(8.3)	81	(3.8)	284	(13.5)	316	(11.6)	366	(9.8)	421	(9.3)	474	(8.4)	520	(10.8)	550	(11.7)
Mangistau region	391	(9.1)	84	(3.6)	255	(12.1)	285	(9.5)	335	(10.0)	388	(9.3)	445	(10.7)	501	(14.6)	535	(13.1)
North-Kazakhstan region	433	(5.4)	83	(2.6)	298	(9.6)	328	(7.3)	378	(5.6)	431	(5.9)	488	(7.5)	540	(8.2)	571	(10.5)
Pavlodar region	438	(5.9)	82	(3.3)	304	(8.9)	332	(9.3)	381	(7.1)	438	(7.1)	493	(7.9)	543	(7.3)	573	(8.8)
South-Kazakhstan region	401	(5.6)	83	(2.8)	267	(10.6)	297	(7.5)	348	(6.6)	401	(5.9)	454	(6.1)	506	(8.4)	539	(10.6)
West-Kazakhstan region	418	(6.5)	82	(2.4)	283	(11.1)	313	(8.3)	362	(8.6)	418	(7.3)	474	(7.4)	524	(8.5)	554	(8.9)
Zhambyl region	456	(6.0)	74	(2.8)	331	(9.9)	360	(8.7)	406	(6.2)	456	(6.5)	507	(6.5)	550	(7.7)	576	(8.1)
Russia																		
*Moscow region**	495	(4.2)	81	(2.0)	360	(7.1)	388	(7.2)	439	(5.4)	496	(4.9)	552	(4.1)	598	(4.6)	625	(6.1)
*Republic of Tatarstan**	475	(3.1)	84	(1.8)	337	(4.3)	367	(3.8)	417	(3.4)	475	(3.4)	533	(3.5)	584	(4.5)	614	(5.6)

* PISA adjudicated region.

Note: See Table I.B1.5 for national data.

StatLink http://dx.doi.org/10.1787/888934029109

Table I.B2.11 [1/2] **Mean score and variation in science performance, by region**

OECD

	Mean score		Standard deviation		Percentiles													
					5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Belgium																		
*Flemish Community**	510	(3.3)	100	(2.1)	334	(6.6)	370	(5.9)	439	(5.3)	518	(4.3)	584	(3.4)	635	(3.4)	663	(3.9)
French Community	485	(2.8)	95	(1.7)	322	(5.5)	354	(5.7)	418	(4.2)	491	(3.6)	554	(3.1)	605	(3.9)	633	(4.0)
German-speaking Community	483	(7.4)	86	(3.2)	335	(11.4)	363	(12.2)	425	(10.1)	487	(8.8)	544	(9.2)	596	(11.9)	621	(9.9)
Canada																		
Alberta	534	(4.4)	96	(2.4)	369	(7.6)	404	(6.3)	468	(5.8)	538	(4.8)	602	(5.0)	654	(6.3)	684	(7.6)
British Columbia	517	(5.4)	101	(2.6)	346	(9.1)	383	(7.5)	446	(5.7)	519	(6.2)	589	(6.6)	647	(6.9)	679	(7.4)
Manitoba	489	(3.7)	95	(1.7)	337	(7.2)	366	(5.6)	423	(5.1)	489	(4.4)	556	(4.8)	612	(4.0)	645	(6.4)
New Brunswick	492	(5.7)	96	(2.8)	336	(9.8)	369	(8.5)	427	(7.0)	493	(6.8)	559	(6.4)	617	(7.6)	650	(10.3)
Newfoundland and Labrador	506	(6.4)	94	(2.7)	354	(11.2)	387	(9.4)	442	(7.2)	505	(7.4)	569	(6.5)	628	(9.6)	663	(10.5)
Nova Scotia	508	(4.7)	94	(2.1)	349	(7.9)	383	(7.2)	444	(6.3)	510	(4.8)	574	(5.1)	629	(6.6)	662	(8.3)
Ontario	519	(4.0)	95	(1.7)	361	(5.8)	395	(4.9)	453	(5.2)	519	(5.1)	587	(4.9)	641	(5.0)	672	(5.5)
Prince Edward Island	502	(8.9)	97	(4.4)	335	(16.5)	369	(16.6)	436	(12.2)	508	(10.7)	571	(10.5)	625	(16.5)	654	(15.7)
Québec	522	(3.7)	91	(1.9)	365	(7.2)	401	(6.0)	461	(4.5)	526	(4.1)	585	(4.3)	635	(4.0)	663	(5.4)
Saskatchewan	501	(3.9)	91	(2.2)	346	(7.7)	382	(6.4)	440	(5.3)	502	(4.3)	564	(4.2)	617	(6.0)	647	(6.9)
Colombia																		
Bogotá	451	(4.9)	84	(2.6)	316	(6.4)	343	(5.9)	391	(5.9)	449	(6.0)	509	(6.4)	560	(7.0)	590	(7.4)
Italy																		
Bolzano	498	(4.2)	84	(1.9)	354	(7.0)	385	(5.8)	440	(5.5)	502	(5.0)	559	(5.0)	605	(6.1)	630	(6.0)
Sardegna	452	(3.9)	83	(1.9)	318	(6.9)	345	(5.4)	393	(4.9)	451	(4.9)	509	(4.8)	560	(5.3)	589	(6.2)
Toscana	475	(4.2)	85	(2.4)	327	(8.3)	359	(7.2)	418	(5.9)	479	(5.5)	536	(4.9)	582	(5.0)	608	(5.5)
Trento	495	(2.1)	84	(1.7)	353	(6.1)	383	(5.0)	437	(2.9)	497	(3.8)	554	(3.7)	605	(4.6)	633	(6.0)
Spain																		
Andalusia	471	(4.4)	90	(2.3)	324	(6.8)	353	(5.1)	406	(5.8)	471	(5.6)	535	(5.8)	589	(5.6)	619	(7.2)
Aragon	493	(5.3)	89	(1.8)	340	(8.7)	373	(6.8)	433	(5.8)	498	(5.8)	557	(5.3)	606	(5.8)	635	(7.4)
Asturias	496	(4.8)	89	(2.2)	348	(7.3)	379	(6.9)	433	(5.4)	500	(5.1)	561	(5.6)	609	(6.2)	636	(8.3)
Balearic Islands	482	(5.2)	84	(1.7)	339	(8.0)	371	(7.5)	425	(5.8)	483	(5.7)	542	(5.7)	591	(6.7)	618	(7.0)
Basque Country	487	(4.2)	85	(1.5)	344	(5.4)	374	(5.2)	429	(5.0)	490	(4.4)	547	(4.6)	595	(5.1)	624	(6.1)
Canary Islands	470	(4.5)	87	(1.7)	326	(8.3)	357	(5.7)	410	(5.3)	470	(4.9)	530	(5.0)	582	(5.8)	612	(6.0)
Cantabria	495	(9.2)	86	(2.3)	351	(10.3)	381	(9.4)	436	(9.4)	498	(9.9)	556	(10.4)	606	(10.2)	633	(11.1)
Castile and León	501	(5.0)	88	(1.9)	348	(8.0)	383	(7.5)	443	(6.9)	506	(5.5)	563	(5.0)	611	(5.8)	639	(5.6)
Castile-La Mancha	484	(6.1)	88	(1.7)	339	(8.0)	369	(7.4)	422	(7.0)	485	(6.5)	549	(6.1)	598	(6.9)	625	(8.1)
Catalonia	489	(4.7)	91	(2.3)	337	(7.8)	368	(8.0)	427	(5.7)	492	(5.5)	554	(5.2)	604	(5.6)	632	(6.0)
Ceuta	415	(6.6)	83	(3.7)	290	(11.0)	314	(10.4)	354	(10.0)	410	(8.7)	470	(10.2)	528	(12.5)	563	(12.2)
Comunidad Valenciana	478	(4.4)	86	(1.9)	336	(7.6)	367	(6.5)	418	(4.5)	479	(5.2)	538	(5.4)	588	(6.0)	615	(6.5)
Extremadura	473	(5.9)	87	(1.7)	331	(9.1)	362	(7.5)	414	(7.1)	473	(6.0)	534	(6.3)	587	(5.8)	617	(6.6)
Galicia	510	(4.0)	89	(2.3)	357	(9.3)	392	(6.5)	452	(4.9)	515	(4.2)	574	(5.0)	622	(6.2)	650	(7.5)
La Rioja	487	(7.9)	90	(2.0)	335	(8.0)	367	(8.1)	425	(7.7)	489	(8.2)	550	(8.7)	600	(8.6)	629	(9.0)
Madrid	487	(3.0)	89	(1.6)	338	(4.5)	369	(4.7)	425	(4.4)	489	(3.4)	550	(3.3)	600	(3.4)	628	(3.6)
Melilla	439	(7.6)	86	(4.5)	301	(18.5)	332	(14.0)	378	(9.2)	435	(9.1)	499	(12.9)	555	(12.4)	584	(20.5)
Murcia	479	(5.7)	93	(2.3)	324	(8.1)	355	(8.0)	413	(6.9)	482	(6.2)	547	(5.9)	598	(7.1)	627	(7.3)
Navarre	492	(6.0)	87	(2.4)	347	(8.0)	377	(7.3)	432	(6.5)	494	(6.7)	554	(6.3)	603	(7.2)	632	(8.9)
United Kingdom																		
England	507	(3.0)	100	(1.6)	340	(5.6)	375	(4.6)	439	(3.8)	509	(3.2)	578	(3.6)	635	(3.8)	667	(4.3)
Northern Ireland	491	(4.6)	92	(2.1)	337	(6.8)	370	(5.7)	428	(5.5)	494	(5.4)	558	(5.9)	609	(6.2)	637	(7.4)
*Scotland**	490	(4.0)	98	(2.9)	332	(7.5)	366	(5.7)	422	(4.8)	490	(5.0)	558	(4.7)	617	(5.9)	650	(5.8)
Wales	488	(3.8)	89	(1.5)	340	(5.5)	371	(5.3)	426	(4.4)	490	(4.5)	552	(4.3)	603	(4.6)	632	(5.3)

* PISA adjudicated region.

Note: See Table I.B1.6 for national data.

StatLink http://dx.doi.org/10.1787/888934029109

Table I.B2.11 [2/2] **Mean score and variation in science performance, by region**

	Mean score		Standard deviation		Percentiles													
					5th		10th		25th		Median (50th)		75th		90th		95th	
	Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
Partners																		
Argentina																		
*CABA**	455	(5.4)	88	(2.1)	308	(6.7)	339	(6.7)	393	(6.8)	456	(6.3)	517	(6.3)	569	(5.4)	598	(6.7)
*Córdoba**	427	(4.6)	86	(2.5)	286	(6.6)	316	(5.8)	366	(5.0)	426	(5.1)	488	(6.3)	542	(6.5)	569	(8.4)
*PBA**	413	(5.5)	89	(2.4)	273	(7.3)	302	(6.5)	350	(6.5)	410	(6.7)	476	(7.1)	531	(6.9)	562	(7.1)
*Tucumán**	391	(5.1)	87	(2.5)	256	(6.2)	283	(6.0)	329	(4.8)	385	(5.8)	449	(7.4)	507	(7.6)	541	(8.4)
Brazil																		
North	384	(6.0)	81	(2.8)	263	(8.0)	285	(7.8)	327	(6.6)	378	(7.4)	435	(7.5)	494	(10.2)	530	(9.9)
Northeast	383	(3.7)	88	(3.1)	254	(5.3)	278	(4.9)	319	(3.5)	372	(4.0)	440	(5.6)	504	(6.5)	542	(9.9)
South	419	(5.9)	90	(3.2)	282	(9.0)	308	(5.8)	354	(6.9)	414	(6.9)	480	(8.6)	542	(8.4)	574	(8.3)
Southeast	414	(3.0)	89	(1.8)	276	(3.7)	301	(3.4)	348	(2.9)	408	(3.6)	475	(4.7)	534	(5.4)	569	(6.8)
Middle-West	415	(8.1)	95	(6.5)	278	(10.6)	302	(9.3)	346	(9.3)	406	(9.0)	473	(9.7)	547	(17.3)	593	(27.3)
Indonesia																		
DI Yogyakarta	439	(5.2)	74	(2.9)	322	(4.7)	346	(3.9)	386	(5.5)	437	(6.1)	491	(6.9)	537	(9.2)	564	(9.6)
DKI Jakarta	428	(6.4)	76	(4.5)	314	(5.7)	335	(4.9)	373	(5.1)	422	(6.6)	477	(9.0)	530	(14.3)	563	(17.3)
Kazakhstan																		
Akmola region	401	(5.0)	77	(2.8)	278	(9.0)	304	(7.1)	348	(6.1)	399	(5.8)	453	(6.8)	501	(6.3)	529	(8.0)
Aktobe region	389	(5.1)	68	(2.5)	286	(8.6)	308	(6.3)	344	(4.9)	384	(4.9)	428	(5.5)	474	(8.8)	509	(10.2)
Almaty	431	(8.4)	84	(4.5)	303	(8.6)	327	(9.1)	370	(7.8)	423	(9.9)	485	(12.1)	547	(15.7)	583	(14.8)
Almaty region	380	(4.6)	66	(2.7)	276	(8.6)	298	(6.5)	336	(5.3)	378	(5.0)	423	(5.1)	465	(8.2)	494	(9.1)
Astana	428	(7.6)	83	(4.4)	301	(11.3)	326	(8.7)	369	(8.3)	423	(8.6)	482	(10.1)	537	(10.9)	570	(12.0)
Atyrau region	361	(5.4)	64	(3.0)	263	(9.4)	282	(8.4)	319	(6.9)	358	(5.6)	399	(5.9)	440	(7.1)	469	(9.8)
East-Kazakhstan region	413	(5.6)	75	(3.4)	297	(9.2)	320	(7.8)	361	(6.3)	408	(6.6)	463	(8.3)	516	(8.4)	543	(7.7)
Karagandy region	428	(7.3)	83	(4.2)	304	(6.8)	328	(6.1)	370	(6.2)	419	(7.2)	481	(10.8)	543	(14.6)	576	(13.3)
Kostanay region	426	(5.9)	75	(2.4)	305	(10.3)	330	(7.7)	373	(7.0)	425	(7.4)	476	(7.1)	526	(6.9)	553	(6.8)
Kyzyl-Orda region	374	(4.8)	60	(2.9)	280	(9.6)	302	(7.8)	335	(6.5)	372	(5.6)	411	(5.1)	450	(6.3)	477	(6.3)
Mangistau region	365	(4.9)	66	(2.6)	267	(6.0)	288	(5.5)	320	(4.4)	359	(5.9)	403	(6.5)	450	(10.1)	484	(11.2)
North-Kazakhstan region	419	(5.2)	76	(2.7)	298	(8.8)	325	(6.4)	368	(5.4)	417	(5.8)	469	(6.6)	517	(8.2)	549	(10.6)
Pavlodar region	413	(6.0)	77	(3.1)	295	(8.6)	317	(6.9)	358	(7.4)	408	(6.6)	465	(8.0)	518	(8.4)	547	(8.9)
South-Kazakhstan region	373	(3.7)	65	(2.3)	271	(7.4)	293	(5.7)	331	(4.7)	372	(4.2)	412	(4.8)	453	(4.8)	482	(7.0)
West-Kazakhstan region	391	(5.1)	71	(2.8)	283	(7.2)	306	(5.5)	342	(5.6)	385	(5.7)	436	(6.6)	484	(7.8)	514	(10.0)
Zhambyl region	397	(4.5)	63	(2.1)	304	(6.2)	322	(6.0)	355	(5.4)	392	(5.4)	434	(5.5)	480	(6.3)	511	(8.0)
Russia																		
*Moscow region**	485	(4.1)	81	(2.3)	353	(6.9)	380	(5.9)	428	(5.8)	485	(4.6)	543	(4.4)	590	(4.2)	616	(5.8)
*Republic of Tatarstan**	464	(2.8)	80	(1.5)	337	(3.4)	362	(3.4)	407	(3.3)	461	(3.3)	517	(3.2)	568	(4.2)	600	(5.3)

* PISA adjudicated region.

Note: See Table I.B1.6 for national data.

StatLink http://dx.doi.org/10.1787/888934029109

Annex B2 **List of tables available on line**

https://doi.org/10.1787/888934029109

ANNEX B3
PISA 2018 system-level indicators

System-level data that are not derived from the PISA 2018 student or school questionnaire are extracted from the OECD's annual publication *Education at a Glance* for those countries and economies that participate in that periodic data collection. For other countries and economies, a special system-level data collection was conducted in collaboration with PISA Governing Board members and National Project Managers.

For further information see: *System-level data collection for PISA 2018: Sources, comments and technical notes.pdf* at www.oecd.org/pisa/.

The following tables are available on line at https://doi.org/10.1787/888934029128.

1	Expenditure	Table B3.1.1	Cumulative expenditure by educational institutions per student aged 6 to 15 (2015)
		Table B3.1.2	Teachers' salaries (2017)
		Table B3.1.3	Teachers' salaries (2017)
		Table B3.1.4	GDP per capita (2015, 2016, 2017, 2018)
2	Time and human resources	Table B3.2.1	Teachers' actual teaching time (2018)
		Table B3.2.2	Intended instruction time in compulsory general education, by age (2018)
		Table B3.2.3	School support staff
3	Education system characteristics	Table B3.3.1	Theoretical starting age and theoretical duration (2015)
		Table B3.3.2	Cut-off birthdate for eligibility to school enrolment and first day of the school year (2018)
		Table B3.3.3	Selecting students for different programmes (2018)
4	Accountability	Table B3.4.1	School inspection at the primary level (2018)
		Table B3.4.2	School inspection at the lower secondary level (2018)
		Table B3.4.3	School inspection at the upper secondary level (2018)
		Table B3.4.4	School board
5	Policies and curriculum	Table B3.5.1	Bullying policies
		Table B3.5.2	Civic education
6	School choice	Table B3.6.1	Freedom for parents to choose a public school for their child(ren) (2018)
		Table B3.6.2	Financial incentives and disincentives for school choice (2018)
		Table B3.6.3	Government regulations that apply to schools at the primary and lower secondary levels (2018)
		Table B3.6.4	Criteria used by public and private schools when assigning and selecting students (2018)
		Table B3.6.5	Expansion of school choice within the public school sector over the past 10 years (2018)
		Table B3.6.6	Government-dependent private schools and their role in providing compulsory education at the primary and lower secondary level (2018)
		Table B3.6.7	Independent private schools and their role in providing compulsory education at the primary and lower secondary level (2018)
		Table B3.6.8	Homeschooling as a legal means of providing compulsory education at the primary and lower secondary level (2018)
		Table B3.6.9	Use of public resources for transporting students (2018)
		Table B3.6.10	Responsibility for informing parents about school choices available to them (2018)
		Table B3.6.11	Availability of school vouchers (or scholarships) (2018)
		Table B3.6.12	Extent to which public funding follows students when they leave for another public or private school (2018)

ANNEX C

Released items from the PISA 2018 computer-based reading assessment

ANNEX C

Released items from the PISA 2018 computer-based reading assessment

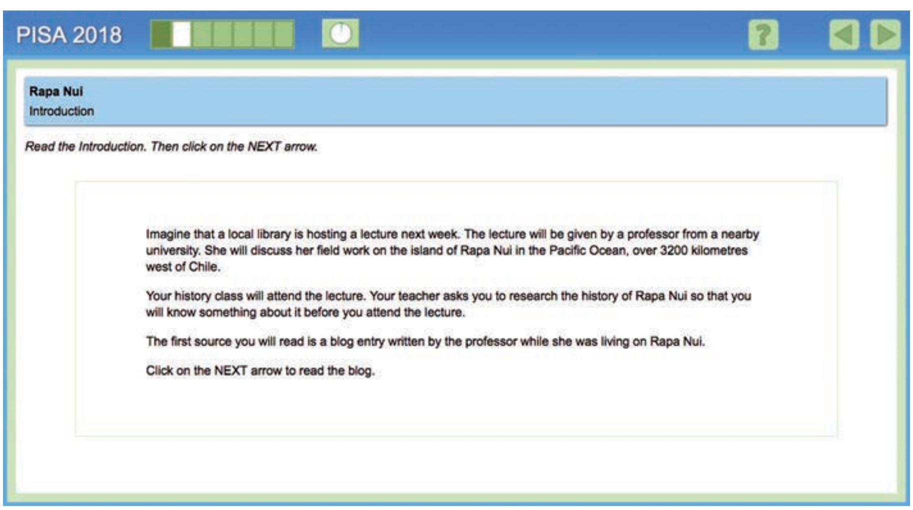
Items from Rapa Nui, the released unit from the PISA 2018 reading assessment, and items used in the assessment of reading fluency

One new unit, *Rapa Nui*, was released from the main survey of the PISA 2018 computer-based reading assessment; the seven items from this unit are presented in this annex. Two other units, *Chicken Forum* and *Cow's Milk*, were tested in the field trial but not used in the PISA 2018 main survey. These units, along with the untested unit *The Galapagos Islands*, are available on line at www.oecd.org/pisa. All four of these units were developed in accordance with the new PISA 2018 reading literacy framework. The annex concludes with sentences that illustrate those used in the reading-fluency assessment.

Screenshots of the interface used in PISA 2018 are shown to give readers an understanding of how students interacted with the assessment and its items. Interactive versions of all of these units are also available at www.oecd.org/pisa.

UNIT CR551: *RAPA NUI*

Rapa Nui *scenario*



In this unit's scenario, the student is preparing to attend a lecture about a professor's field work, which was conducted on the island of Rapa Nui. The situation is classified as educational because it represents a student conducting background research on Rapa Nui in preparation to attend a lecture.

Rapa Nui is a multiple-source unit. It consists of three texts: a webpage from the professor's blog, a book review, and a news article from an online science magazine. The blog is classified as a multiple-source text; dynamic (the webpage contains active links to the other texts in the unit); continuous; and narrative. The blog post is an example of a multiple-source text because the comment section at the bottom of the blog page represents different authors. Both the book review and the news article are classified as single text; static; continuous; and argumentative.

Initially, the student is provided with the blog post only. Several questions are presented that focus only on the content of this blog. Once those questions have been answered, the student receives the second text – the book review. After reading the book review, the student responds to a question that focuses solely on its content. The student then receives the third text – the article from the online science magazine. The student sees questions that focus only on the article. After that, the student is given items that require integrating the information from all sources.

This model was used for several of the multiple-text units in the new material developed for reading literacy. This approach was chosen because it allows the student first to demonstrate proficiency on questions that are related to one text and then to demonstrate the ability to handle information from multiple texts. This is an important design feature because there may be readers who can succeed with information when it is presented in a single text and even integrate information within one text, but who struggle when asked to integrate across multiple texts. Thus, this design allows students with varying levels of ability to demonstrate proficiency on at least some elements of the unit.

The "Rapa Nui" unit was intended to be of moderate to high difficulty. The three texts result in a larger amount of information to work through within the unit compared to a single-text unit. In addition, the student needs to consider the way the texts are related to one another, requiring him or her to recognise whether the texts corroborate each other or whether they differ in their stances. This kind of cognitive engagement with the material and the unit overall is expected to require more effort than a unit that presents all the information within one text.

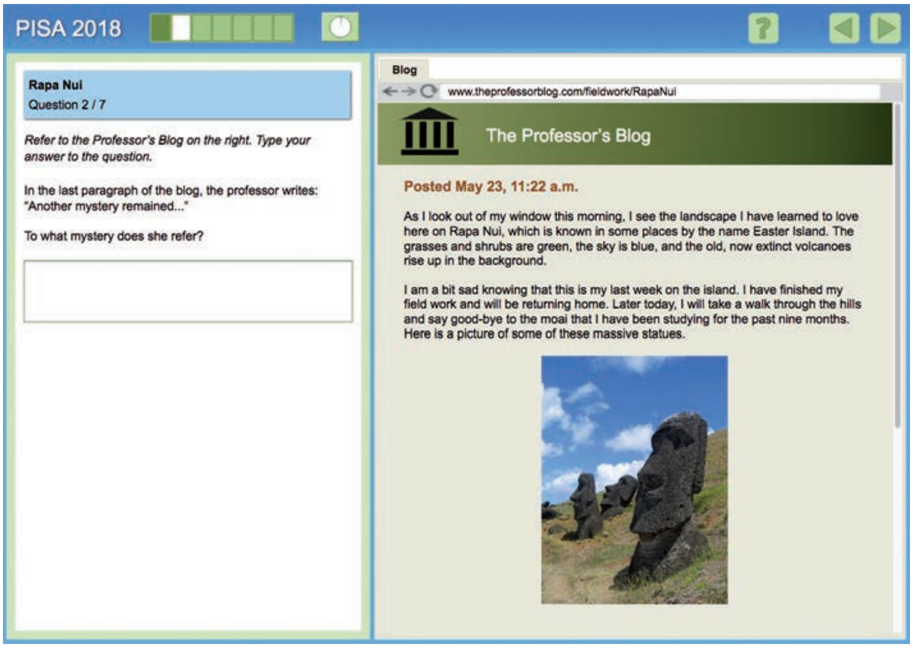
Please note that the screenshot provided for released item #1 shows the full text of the blog for the purposes of this report. The student had to scroll to see the full text in the programmed version, which was programmed uniformly across language versions so that all students would have to scroll to see the full text.

***Rapa Nui** released item #1*

In this item, the student must locate the correct information within the blog post. The difficulty of the item likely stems from the existence of other time-related information within the blog, i.e. the date it was posted and the time period in which the first mystery of the moai was solved (the 1990s). Here, the correct answer is (B) Nine months ago.

Item number	CR551Q01
Cognitive process	Accessing and retrieving information within a piece of text
Response format	Simple multiple choice – Computer scored
Difficulty	559 – Level 4
Source type	Single source

Rapa Nui *released item #2*



In this item, the student must understand that the second mystery mentioned in the blog post: what happened to the large trees that once grew on Rapa Nui and were used to move the moai? This is an open response/human coded item, and the coding guide used in the main survey is provided below. For this item, the student could provide a direct quotation from the blog ("What happened to these plants and large trees that had been used to move the moai?") or an accurate paraphrase. This item was coded with high reliability in the main survey.

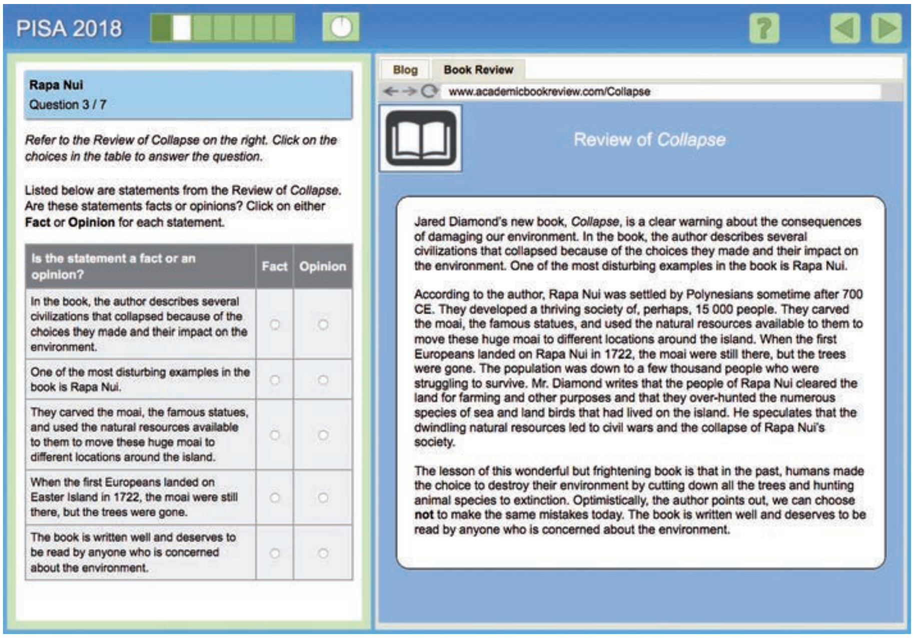
Item number	CR551Q05
Cognitive process	Representing literal meaning
Response format	Open response – Human coded
Difficulty	513 – Level 3
Source type	Single source

For full credit, responses must refer to the disappearance of the materials used to move the statues (moai).

- What happened to these plants and large trees that had been used to move the moai? *[Direct quotation]*
- There are no large trees left that could have moved the moai.
- There are grasses, shrubs and some small trees, but no trees large enough to move the large statues.
- Where are the large trees? *[Minimal]*
- Where are the plants? *[Minimal]*
- What happened to the resources that were needed to transport the statues?
- She was referring to what moved the Moai because when she looked around there were no big trees or plants. She is also wondering what happened to them. *[Although this response begins by referring to the wrong mystery, it contains the correct elements.]*

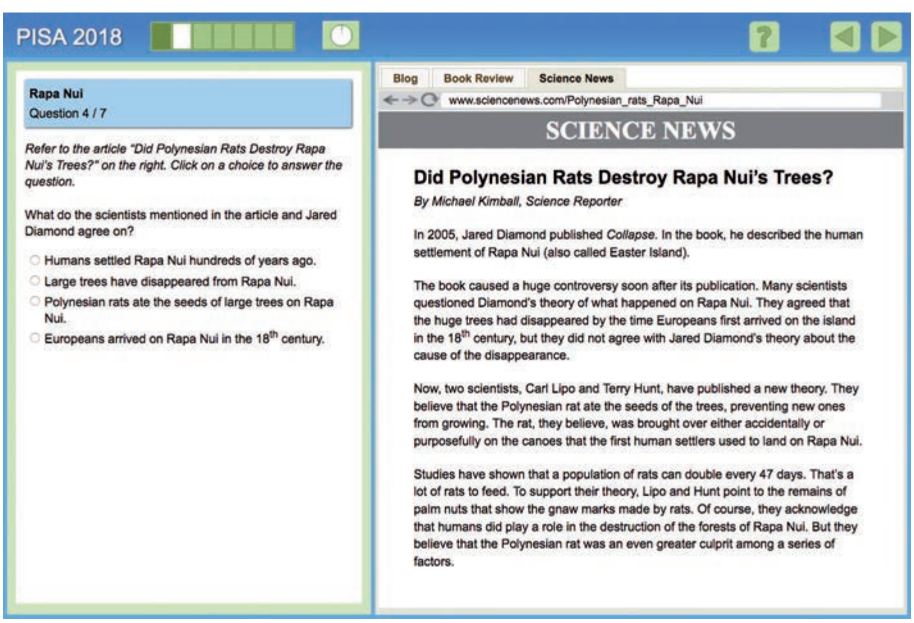
Rapa Nui *released item #3*

For this item, the student is presented with the second text in the unit, a book review of *Collapse*, which was referenced in the blog post. The student must complete a table by selecting "Fact" or "Opinion" for each row. The question asks the student to identify whether each statement from the book review is a fact or an opinion. The student must first understand the literal meaning of each statement and then decide if the content was factual or represented the perspective of the author of the review. In this way, the student must focus on the content and how it is presented rather than just the meaning. To receive full credit for this item, the student was required to get all 5 rows correct. For partial credit, students were required to get 4 out of the 5 rows correct. If students got fewer than 4 rows correct, they received no credit. The correct answers are: Fact, Opinion, Fact, Fact, Opinion.



Item number	CR551Q06
Cognitive process	Reflecting on content and form
Response format	Complex multiple choice – Computer scored
Difficulty	For full credit, 654 – Level 5; for partial credit, 528 – Level 3
Source type	Single source

***Rapa Nui** released item #4*

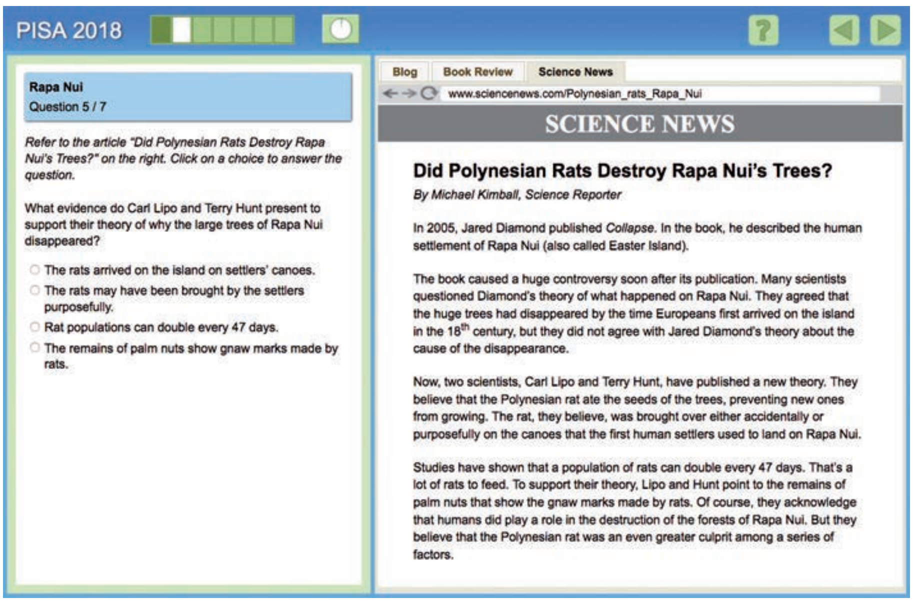


For this item, the student is presented with the third text in the unit – an article from an online science magazine. Note that at this point in the unit all three texts are available to the student using a tab structure; the student can click on any tab to toggle back and forth between the texts. The item itself remains fixed on the left side of the screen during any toggling action. In this item, the student is required to locate the section of the article that contains the reference to the scientists and Jared Diamond (paragraph 2) and identify the sentence that contains the information agreed upon. While texts are available to the student, this item is not classified with a cognitive process that reflects the use of multiple sources. This is because the student can find the answer within this text, and the item instructions on the upper left corner instruct the student to refer to this article only.

Thus, the support from the item instructions eliminates the need to consider the other sources. The difficulty of this item is likely driven by the existence of plausible (but incorrect) distracting information within the paragraph with respect to human settlement. Here, the correct answer is (B) Large trees have disappeared from Rapa Nui.

Item number	CR551Q08
Cognitive process	Accessing and retrieving information within a piece of text
Response format	Simple multiple choice – Computer scored
Difficulty	634 – Level 5
Source type	Single source

***Rapa Nui** released item #5*

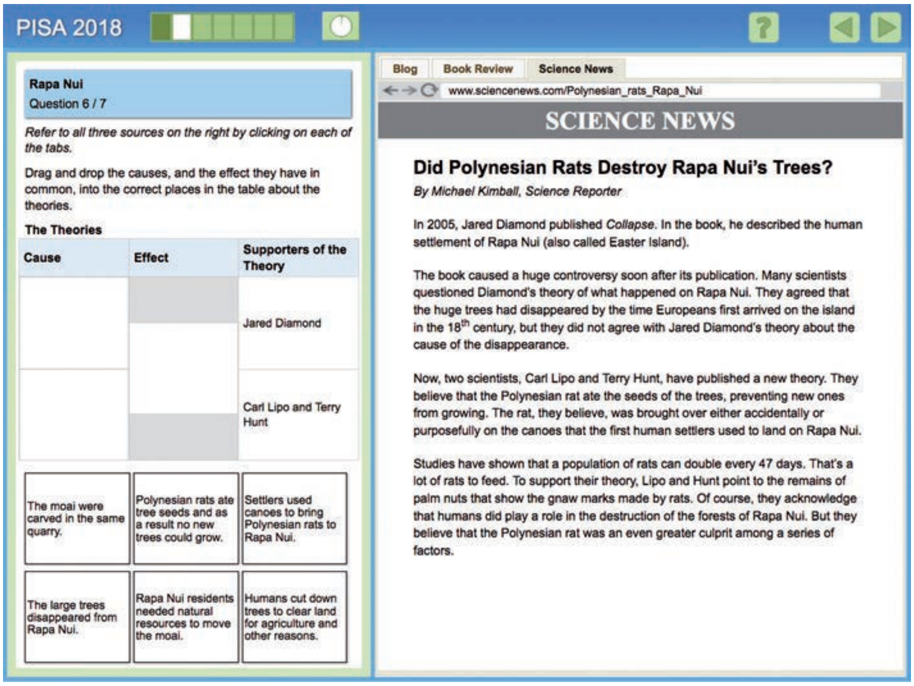


In this item, the student is required to understand what information in the text supports, or corroborates, the theory put forward by the scientists. The correct answer is (D) The remains of palm nuts show gnaw marks made by rats. Here, the student must go beyond an understanding of the text and identify which element of the text can be used as evidence to support a claim. All other items classified as detect and handle conflict require detecting a conflict between two sources or recognising that the information is in two or more sources and is corroborated. However, in discussing this item prior to the field trial, the experts felt that the act of identifying which piece of information supports the theory proposed by Carl Lipo and Terry Hunt was most appropriately identified by the cognitive process of detect and handle conflict. Furthermore, while this item could have been classified as requiring only a single source in order to be solved, the requirement for the student to first consider the theory proposed by Lipo and Hunt and then to determine which piece of evidence supports this theory is akin to working with multiple sources.

Item number	CR551Q09
Cognitive process	Detecting and handling conflict
Response format	Simple multiple choice – Computer scored
Difficulty	597 – Level 4
Source type	Multiple source

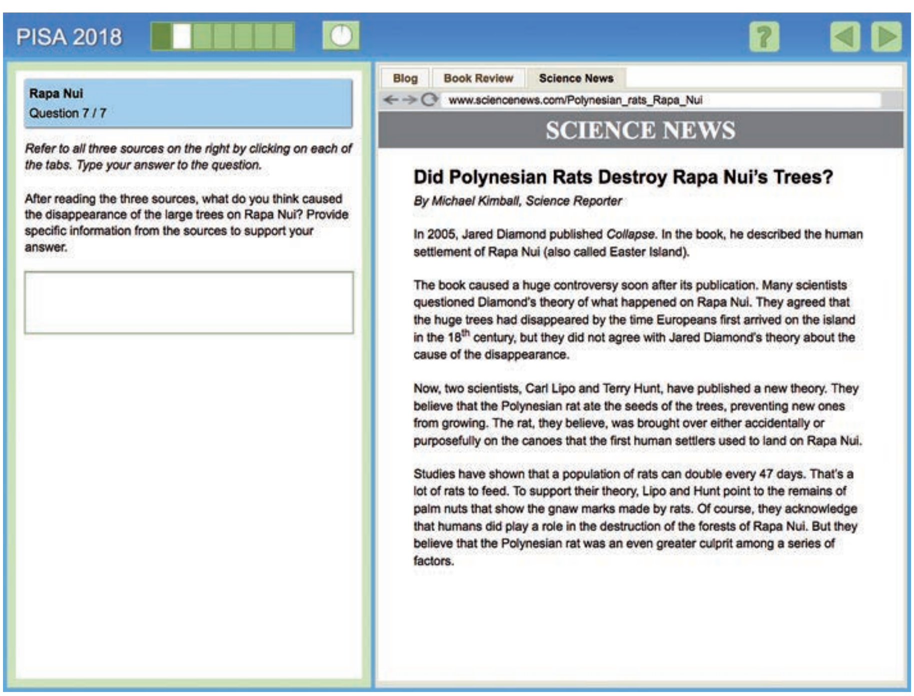
***Rapa Nui** released item #6*

In this item, students must integrate information across the texts with respect to the differing theories put forward by Jared Diamond on the one hand and Carl Lipo and Terry Hunt on the other. The student must identify the shared effect (the disappearance of the large trees) by rejecting information presented in the blog post about where the moai were carved (in the same quarry). Further, the student must understand what each scientist believes is the cause of the disappearance. To receive credit for this item, the student was required to get all three answers correct. The correct answers are: Cause (Jared Diamond) – Humans cut down trees to clear land for agriculture and other reasons. Cause (Carl Lipo and Terry hunt) – Polynesian rats ate tree seeds and as a result no new trees could grow. Effect (shared) – The large trees disappeared from Rapa Nui.



Item number	CR551Q10
Cognitive process	Integrating and generating inferences across multiple sources
Response format	Complex multiple choice – Computer scored
Difficulty	665 – Level 5
Source type	Multiple source

Rapa Nui *released item #7*



In this item, the student must integrate information from across the texts and decide which theory to support. In this way, the student must understand the theories – and that they are at odds with one another – and must present a response that contains support from the texts. To receive credit, a student could choose to support either theory or could choose neither theory as long as the explanation is focused on the need for additional research. This is an open response/human coded item, and the coding guide used in the main survey is provided below. This item was coded with high reliability in the main survey.

Item number	CR551Q11
Cognitive process	Detecting and handling conflict
Response format	Open response – Human coded
Difficulty	588 – Level 4
Source type	Multiple source

For full credit, at least one of the following descriptions had to be included:

1. The people cut down or used the trees (to move the moai and/or cleared the land for agriculture).
2. The rats ate the seeds of the trees (so new trees could not grow).
3. It is not possible to say exactly what happened to the large trees until further research is conducted.

Sample responses that would receive full credit include:

- I think the trees disappeared because people cut too many of them down to move the moai. *[1]*
- People cleared the land for agriculture. *[1]*
- Trees were used to move moai. *[1]*
- People cut the trees down. *[1]*
- It was the people's fault because they wanted to move the moai. *[1 – this response doesn't explicitly refer to cutting down the trees, but it is acceptable because they refer to people and one reason they cut down the trees (to move the moai)]*
- *People's fault. They destroyed the environment. [1 – this response doesn't explicitly refer to cutting down the trees, but it is an acceptable way of summarizing the results of cutting down the trees.]*
- I think the rats probably caused the most damage by eating the seeds of the trees. *[2]*
- The rats ate the seeds. *[2]*
- There is no proof that either one is correct, so we have to wait until there is more information. *[3]*
- Both. The people cut down the big trees for farming, and then the rats ate the tree seeds! *[1 and 2]*

READING FLUENCY

In PISA 2018, the Reading Expert Group recommended including a measure of reading fluency to better assess and understand the reading skills of students in the lower proficiency levels. PISA defines reading fluency as the ease and efficiency with which one can read and understand a piece of text. Reading fluently requires that one can recognize words within a text accurately and automatically and can then parse and process the words into a coherent whole in order to comprehend the overall meaning of the text. When these processes are done efficiently, students' cognitive resources are available for higher-level comprehension tasks, allowing students to engage with texts more deeply.

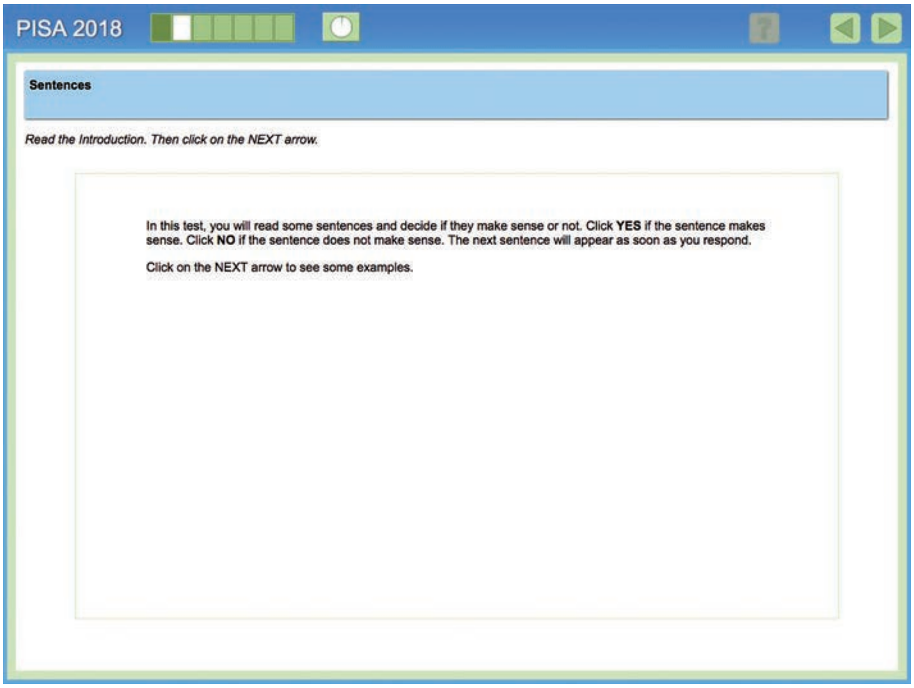
In the PISA 2018 assessment of reading fluency, students were given three minutes to evaluate the sensibility of as many sentences as they could (i.e. Does the sentence make sense – Yes or No). The number of sentences was restricted to 21 or 22 sentences per student so that most students would be able to complete the task within the three minutes. Students were not cut off in the middle of an item or notified that they did not complete all the sentences. Instead, if a student reached the three minutes while viewing a sentence, the task ended after they completed that sentence's sensibility judgment. This was done so that students would maintain motivation for the remaining sections of the PISA assessment.

Items in this task were the easiest items in the reading-literacy assessment in PISA 2018. Difficulty information is not provided in this report for the practice items because data for these items was not analysed. However, in the assessment of reading fluency, the items fell into proficiency Level 1c and Level 1b. One item was in Level 1a. Items that did not make sense and required a "No" response were more difficult than items that made sense and required a "Yes" response.

The introduction to and practice items for the reading-fluency task are provided below along with an explanation for how students were oriented to the task.

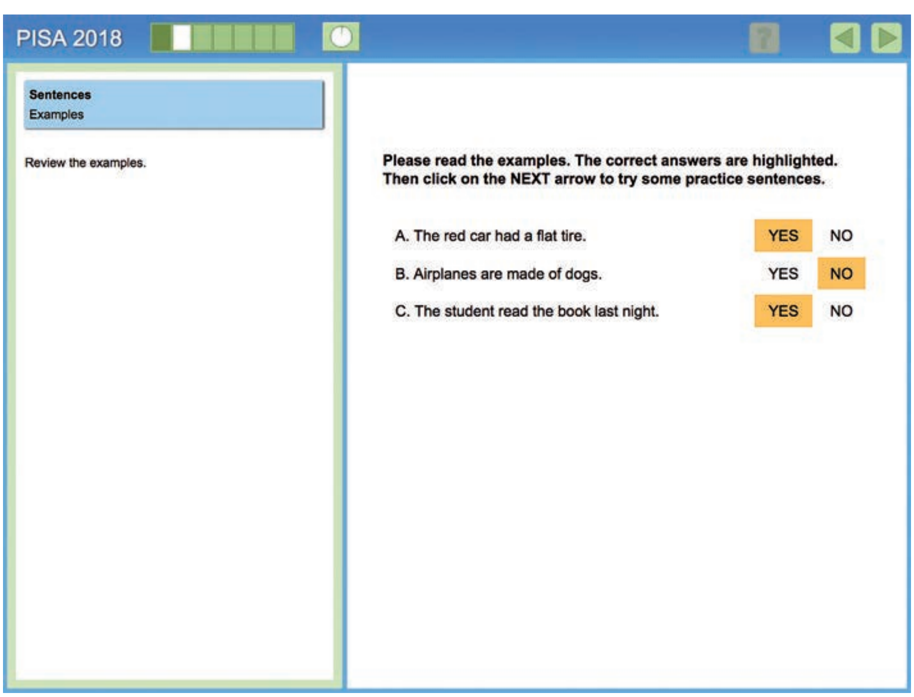
Reading fluency: *Introduction*

In this introduction, students are given the basic instructions for what they will do in the fluency task. Students are notified that the next sentence will appear as soon as they respond so that they are prepared for this style of presentation.



Reading fluency: *Static examples*

Students are given a set of static examples so that the sensibility judgements are understood prior to interacting with dynamic practice items. Here, three example sentences are provided, two that make sense (a Yes response is correct) and one that does not make sense (a No response is correct).

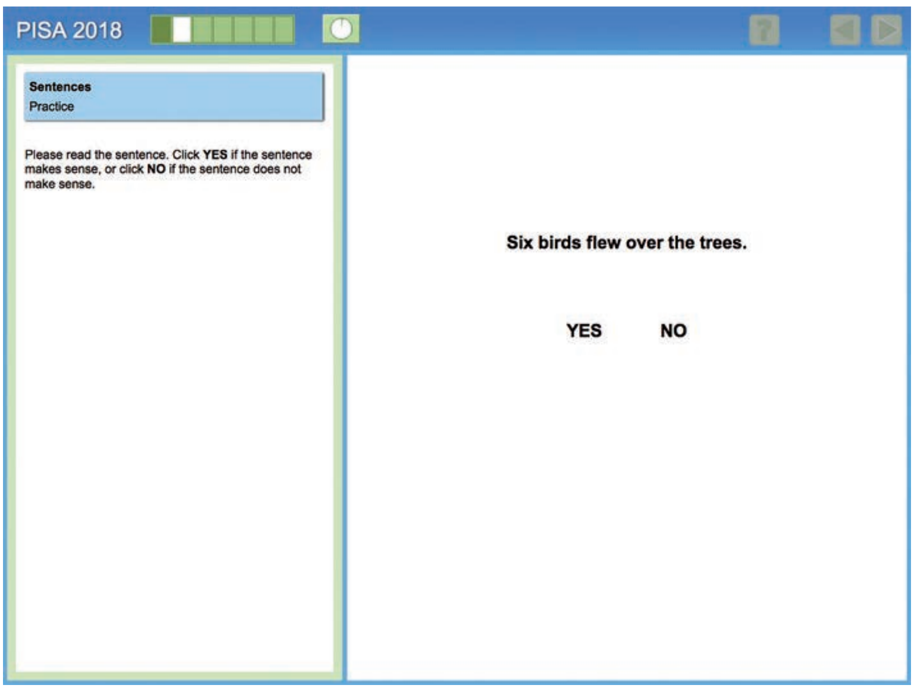


Reading fluency: *Dynamic practice*

The next three images show three dynamic-practice items. Students complete these dynamic-practice items prior to receiving the first fluency item so that they understand the response mode for the item. For each example, as soon as the student clicks on "Yes" or "No", the next item appears.

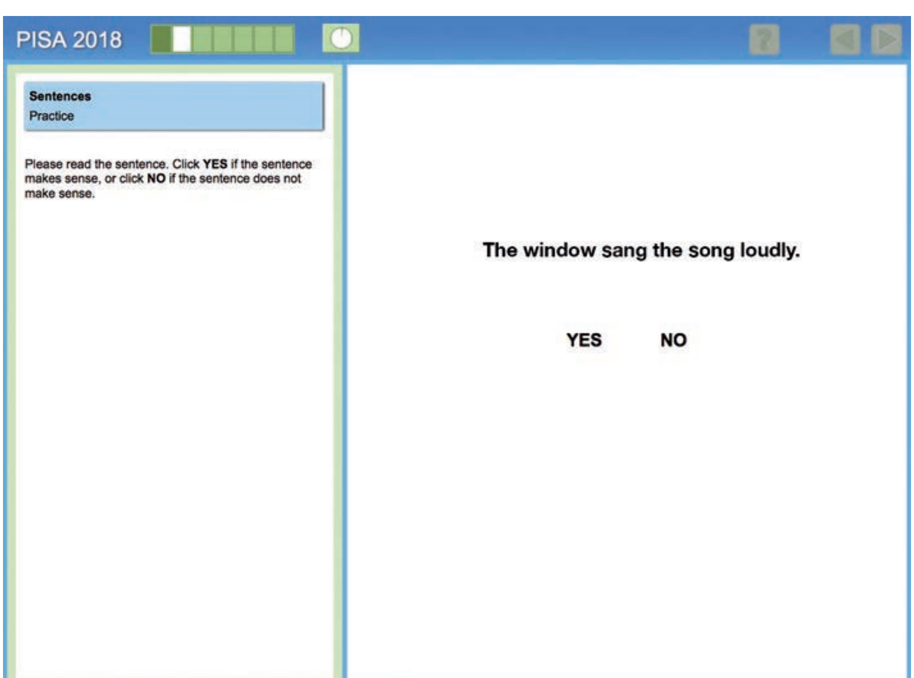
Reading fluency: *Dynamic-practice item #1*

Here, the correct answer is "Yes".



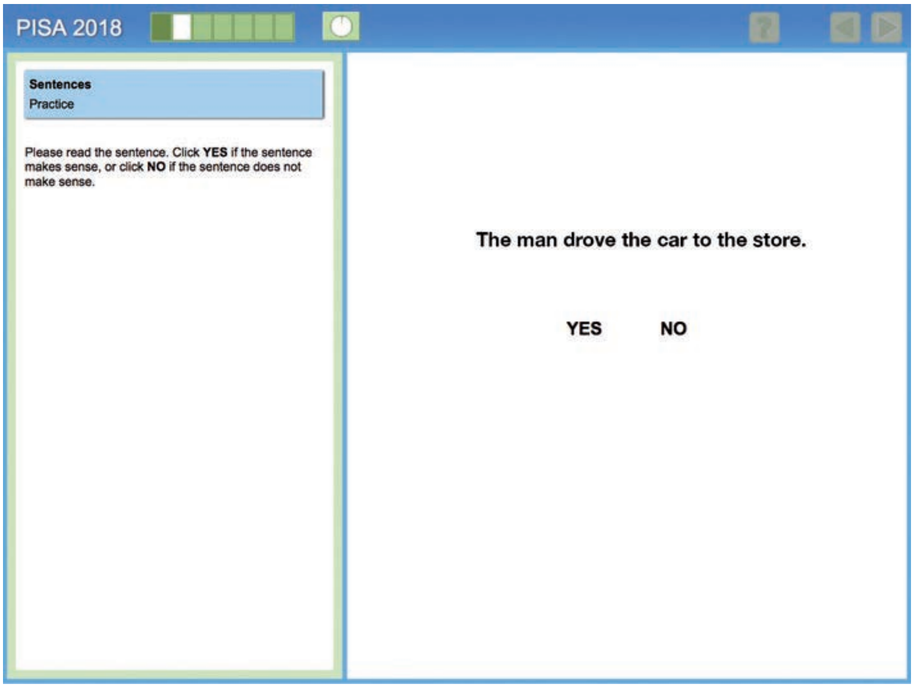
Reading fluency: *Dynamic-practice item #2*

Here, the correct answer is "No".



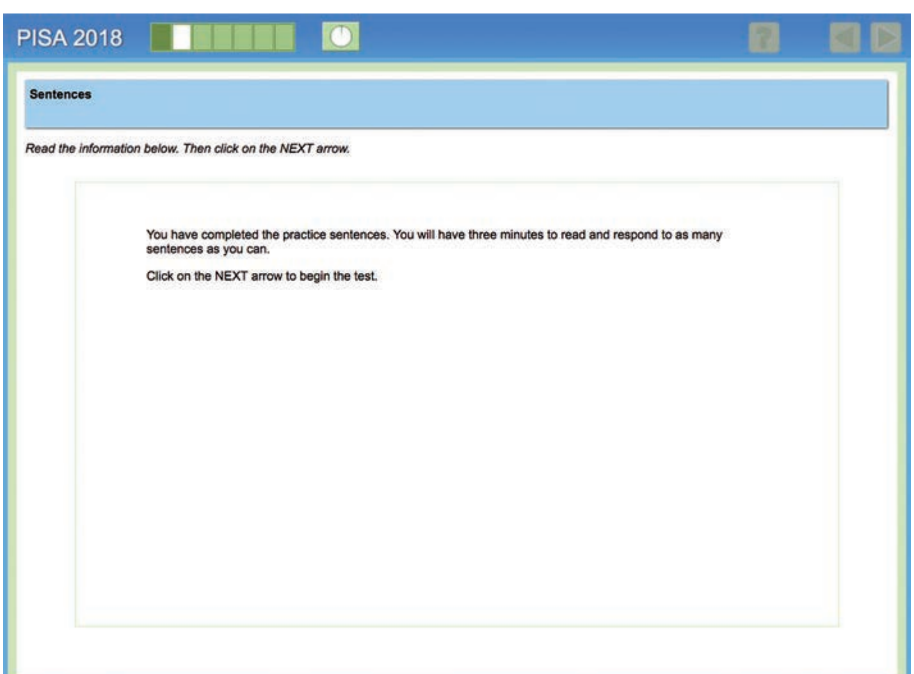
Reading fluency: *Dynamic-practice item #3*

Here, the correct answer is "Yes".



Reading fluency: *End of practice*

Students are told that they have completed the practice sentences. They are also given the time limit for the task – three minutes – and they are told to complete as many sentences as they can within the time limit. Once the student clicks on the NEXT arrow, the task begins and is carried out in the same way as the dynamic-practice items. Once students have completed the task, they are notified that the first section of the test is complete and the answers have been saved.



PISA

ANNEX D

Snapshot of trends in reading, mathematics and science performance

ANNEX D
Snapshot of trends in reading, mathematics and science performance

Snapshot of performance trends in ALBANIA

Mean performance	Reading	Mathematics	Science
PISA 2000	349*		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	385*	377*	391*
PISA 2012	394*	394*	397*
PISA 2015	405	413*	427*
PISA 2018	405	437	417
Average 3-year trend in mean performance	+10.5*	+19.8*	+10.7*
Short-term change in mean performance (2015 to 2018)	+0.2	+24.1*	-10.5*
Overall performance trajectory	positive, but flattening (less positive over more recent years)	improving	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.2	+1.5*	+0.1
Percentage-point change in low-achieving students (below Level 2)	-4.4	-18.3*	-10.3*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+7.9*	+16.7*	+4.1
Average trend amongst the lowest-achieving students (10th percentile)	+14.4*	+24.0*	+19.7*
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Albania, mean performance improved, from initially low levels, across all three subjects (reading, mathematics and science). In all three subjects, improvements at the bottom of the performance distribution outpaced improvements observed at the top, resulting in narrowing performance gaps between the highest- and lowest-achieving students. Improvements in mean performance were particularly rapid in mathematics (about 20 points, on average, per 3-year period). The proportion of students who scored below Level 2 in mathematics (low-achieving students) shrank by 18 percentage points between 2012 and 2018.

Improvements in performance in Albania were even more remarkable when considering that enrolment rates of 15-year-olds in grade 7 and above increased between 2009 and 2018 (Table I.A2.2).

Snapshot of performance trends in ARGENTINA

Mean performance	Reading	Mathematics	Science
PISA 2000	418		
PISA 2003	m	m	
PISA 2006	374*	381	391
PISA 2009	398	388	401
PISA 2012	396	388	406
PISA 2015	m	m	m
PISA 2018	402	379	404
Average 3-year trend in mean performance	-1.2	-1.0	+3.0
Short-term change in mean performance (2015 to 2018)	m	m	m
Overall performance trajectory	U-shaped (more positive over more recent years)	stable	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.3	+0.1	+0.0
Percentage-point change in low-achieving students (below Level 2)	+0.5	+2.5	-2.8
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-4.3*	-5.6*	-0.0
Average trend amongst the lowest-achieving students (10th percentile)	+4.4	+5.3*	+8.3*
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean mathematics and science performance remained stable in Argentina over the 2006-2018 period. In reading, performance improved over this period after an initial decline between 2001 and 2006.

The gap between the highest- and lowest-achieving students narrowed in all three subjects. This means that in reading, mathematics and science, trends were significantly more positive at the bottom of the performance distribution (with the 10th percentile moving up by more than 5 points per 3-year period in mathematics and science) than at the top of the performance distribution (with the 90th percentile moving down by more than 4 points per 3-year period in reading and mathematics).

PISA 2015 results for Argentina cannot be compared to results from previous years or to results from 2018 due to the use of an incomplete sampling frame. Indeed, PISA 2015 results represented only 55% of the country's population of 15-year-olds, compared to about 80% in PISA 2006, 2012 and 2018.

Snapshot of performance trends in AUSTRALIA

Mean performance	Reading	Mathematics	Science
PISA 2000	528*		
PISA 2003	525*	524*	
PISA 2006	513	520*	527*
PISA 2009	515*	514*	527*
PISA 2012	512*	504*	521*
PISA 2015	503	494	510*
PISA 2018	503	491	503
Average 3-year trend in mean performance	-4.4*	-7.2*	-6.5*
Short-term change in mean performance (2015 to 2018)	-0.3	-2.5	-7.0*
Overall performance trajectory	steadily negative	steadily negative	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.3	-4.3*	-5.1*
Percentage-point change in low-achieving students (below Level 2)	+5.4*	+2.8	+6.0*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-2.4	-6.9*	-6.0*
Average trend amongst the lowest-achieving students (10th percentile)	-6.2*	-7.1*	-7.6*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in Australia has been steadily declining in reading (between 2000 and 2018) and in mathematics (between 2003 and 2018), from initially high levels of performance; it has been declining in science too, at least since 2012. In reading, more rapid declines were observed amongst the country's lowest-achieving students. In mathematics and science, performance declined to a similar extent at the top and at the bottom of the performance distribution, as well as on average.

The proportion of top-performing students (scoring at Level 5 or 6) remained stable in reading (between 2009 and 2018), but decreased in mathematics (between 2012 and 2018) and in science (between 2006 and 2018). Meanwhile, the proportion of low-achieving students (scoring below Level 2) increased in all subjects.

Snapshot of performance trends in AUSTRIA

Mean performance	Reading	Mathematics	Science
PISA 2000	492		
PISA 2003	491	506	
PISA 2006	490	505	511*
PISA 2009	m	m	m
PISA 2012	490	506	506*
PISA 2015	485	497	495
PISA 2018	484	499	490
Average 3-year trend in mean performance	-1.3	-1.7	-5.5*
Short-term change in mean performance (2015 to 2018)	-0.5	+2.2	-5.3
Overall performance trajectory	flat	flat	declining
Proficiency levels	**Reading (2012 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.9*	-1.7	-3.7*
Percentage-point change in low-achieving students (below Level 2)	+4.1*	+2.4	+5.5*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-1.3	-2.3	-5.1*
Average trend amongst the lowest-achieving students (10th percentile)	-0.9	-1.7	-4.7*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Austria's mean performance in reading and mathematics remained stable, around a flat trend line, throughout the country's participation in PISA. In science, performance has been declining since 2006; similar declines were observed amongst the country's highest-achieving and lowest-achieving students. In PISA 2018, the proportion of top-performing students in science (students scoring at Level 5 or 6) was almost 4 percentage points smaller than in 2006.

Snapshot of performance trends in BELGIUM

Mean performance	Reading	Mathematics	Science
PISA 2000	507*		
PISA 2003	507	529*	
PISA 2006	501	520*	510*
PISA 2009	506*	515	507
PISA 2012	509*	515	505
PISA 2015	499	507	502
PISA 2018	493	508	499
Average 3-year trend in mean performance	-1.8	-4.1*	-2.7*
Short-term change in mean performance (2015 to 2018)	-5.7	+1.1	-3.2
Overall performance trajectory	hump-shaped (more negative over more recent years)	negative, but flattening (less negative over more recent years)	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-1.6	-3.8*	-2.0*
Percentage-point change in low-achieving students (below Level 2)	+3.5*	+0.7	+3.0
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-1.8	-6.8*	-2.5*
Average trend amongst the lowest-achieving students (10th percentile)	+1.1	-1.1	-2.2
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In all three subjects, Belgium's mean performance in PISA 2018 was not significantly different from that observed in 2015. When considering a longer period, the overall trajectory is negative in mathematics and science, and declining, at least since 2012, in reading too.

The decline in mean performance in mathematics, most of which occurred in the earlier period, was mostly the result of declines amongst the highest-achieving students. The 90th percentile of the mathematics performance distribution, i.e. the level above which only 10% of students scored, moved down by about 7 points per 3-year period between 2003 and 2018.

Snapshot of performance trends in BRAZIL

Mean performance	Reading	Mathematics	Science
PISA 2000	396*		
PISA 2003	403	356*	
PISA 2006	393*	370*	390*
PISA 2009	412	386	405
PISA 2012	407	389	402
PISA 2015	407	377	401
PISA 2018	413	384	404
Average 3-year trend in mean performance	+2.6	+4.6*	+2.2
Short-term change in mean performance (2015 to 2018)	+5.5	+6.5	+2.9
Overall performance trajectory	flat	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.5	+0.2	+0.2
Percentage-point change in low-achieving students (below Level 2)	+0.4	-0.2	-5.6*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.0*	+2.6	+4.0*
Average trend amongst the lowest-achieving students (10th percentile)	+2.6	+7.4*	+1.2
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Brazil, mean performance in mathematics improved over the 2003-2018 period, but most of that improvement was in the early cycles. After 2009, in mathematics, as in reading and science, mean performance appeared to fluctuate around a flat trend.

The positive early trends (2000-2012) were observed over a period of rapid expansion of secondary education. Between 2003 and 2012, Brazil added more than 500 000 students to the total population of 15-year-olds eligible to participate in PISA. The proportion of 15-year-olds who were covered by PISA samples increased from about 55% in 2003 to 70% in 2012. It is likely that this expansion in education opportunities dampened an even more positive underlying trend in student performance. Indeed, a simulation that assumes that the highest-scoring 25% of 15-year-olds were eligible to take the test in any given year shows a positive trend amongst this population not only in mathematics (2003-2018), but also in science (2006-2018) (Figure I.9.5).

Snapshot of performance trends in BULGARIA

Mean performance	Reading	Mathematics	Science
PISA 2000	430		
PISA 2003	m	m	
PISA 2006	402	413*	434
PISA 2009	429	428	439
PISA 2012	436*	439	446*
PISA 2015	432	441	446*
PISA 2018	420	436	424
Average 3-year trend in mean performance	+0.8	+5.9*	-1.4
Short-term change in mean performance (2015 to 2018)	-11.9	-5.1	-21.7*
Overall performance trajectory	flat	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.4	+0.2	-1.5*
Percentage-point change in low-achieving students (below Level 2)	+6.1	+0.7	+3.9
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+1.8	+5.4*	-4.6*
Average trend amongst the lowest-achieving students (10th percentile)	+0.9	+6.2*	+2.0
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Bulgaria, mean performance in reading remained stable, around a flat trend line, throughout the country's participation in PISA (2001-2018). In mathematics, performance improved between 2006 and 2018, but the improvement was concentrated in the early years (2006-2012). In science, performance in 2018 fell below the level observed in 2012 and 2015. The drop in mean science performance between PISA 2015 and PISA 2018 is one of the largest observed over this (short) period amongst all PISA-participating countries and economies.

Snapshot of performance trends in CANADA

Mean performance	Reading	Mathematics	Science
PISA 2000	534*		
PISA 2003	528	532*	
PISA 2006	527	527*	534*
PISA 2009	524	527*	529*
PISA 2012	523	518	525
PISA 2015	527	516	528*
PISA 2018	520	512	518
Average 3-year trend in mean performance	-1.7	-4.1*	-3.4*
Short-term change in mean performance (2015 to 2018)	-6.6	-3.6	-9.7*
Overall performance trajectory	flat	steadily negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.2	-1.1	-3.1*
Percentage-point change in low-achieving students (below Level 2)	+3.5*	+2.4	+3.4*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-0.3	-2.9*	-2.0
Average trend amongst the lowest-achieving students (10th percentile)	-2.8*	-5.5*	-4.3*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Canada, performance declined in mathematics (since 2003) and in science (since 2006) by about 10 score points or more per decade (4.1 score points per 3-year period in mathematics, and 3.4 score points per 3-year period in science). In reading, no significant overall direction of the trend could be determined, and performance remained at least 20 points above the OECD average performance in every PISA year. However, the share of low-achieving students increased between 2009 and 2018 by 3.5 percentage points and, as is observed in mathematics too, more rapid declines were observed amongst the lowest-achieving students than amongst the highest-achieving students, resulting in a widening of performance gaps.

Snapshot of performance trends in CHILE

Mean performance	Reading	Mathematics	Science
PISA 2000	410*		
PISA 2003	m	m	
PISA 2006	442	411	438
PISA 2009	449	421	447
PISA 2012	441*	423	445
PISA 2015	459	423	447
PISA 2018	452	417	444
Average 3-year trend in mean performance	+7.1*	+1.4	+1.1
Short-term change in mean performance (2015 to 2018)	-6.3	-5.3	-3.4
Overall performance trajectory	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.3*	-0.4	-0.9*
Percentage-point change in low-achieving students (below Level 2)	+1.2	+0.4	-4.4
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+6.2*	+0.9	-0.7
Average trend amongst the lowest-achieving students (10th percentile)	+8.1*	+0.9	+1.9
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Reading performance in Chile improved since the country's first participation in PISA (in 2001). However, most of that improvement occurred in the early period. Between 2009 and 2018, no significant trends in performance were observed in any subject.

Despite stable overall performance, the proportion of students performing at Level 5 or above (top performers) in reading grew between 2009 and 2018 (+1.3 percentage points) and shrank in science between 2006 and 2018 (-0.9 of a percentage point).

Snapshot of performance trends in COLOMBIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	385*	370*	388*
PISA 2009	413	381	402
PISA 2012	403	376*	399*
PISA 2015	425*	390	416
PISA 2018	412	391	413
Average 3-year trend in mean performance	+6.6*	+5.1*	+6.4*
Short-term change in mean performance (2015 to 2018)	-12.6*	+1.3	-2.4
Overall performance trajectory	positive, but flattening (less positive over more recent years)	steadily positive	steadily positive
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.4	+0.2	+0.3*
Percentage-point change in low-achieving students (below Level 2)	+2.8	-8.4*	-9.8*
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.5*	+4.9*	+7.3*
Average trend amongst the lowest-achieving students (10th percentile)	+12.0*	+7.2*	+8.0*
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

While Colombia's performance in reading in PISA 2018 was below that observed in 2015, when considering a longer period, mean performance improved in all subjects – including reading – since the country first participated in PISA in 2006.

Snapshot of performance trends in COSTA RICA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	443*	409	430*
PISA 2012	441*	407	429*
PISA 2015	427	400	420
PISA 2018	426	402	416
Average 3-year trend in mean performance	-6.8*	-3.0	-6.1*
Short-term change in mean performance (2015 to 2018)	-1.0	+2.1	-4.0
Overall performance trajectory	declining	stable	declining
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.2	-0.2	-0.2
Percentage-point change in low-achieving students (below Level 2)	+9.3*	+0.1	+8.8*
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-3.7	-1.5	-5.5
Average trend amongst the lowest-achieving students (10th percentile)	-7.6*	-5.0*	-5.2*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Costa Rica first participated in PISA in 2010. While mean performance in mathematics remained stable over the 2010-2018 period, it declined in both reading and science. More specifically, while performance in reading and science was similar between 2009 and 2012, it declined in 2015 and stayed at roughly the same level in 2018. The decline in performance was most acute amongst the lowest-achieving students. The average trend amongst these students was negative and significant in all three subjects (reading, mathematics and science).

However, these decreases in performance took place in the context of an increase in the coverage of the 15-year-old population in Costa Rica, from between 50% and 53% in 2010 and 2012, respectively, to 63% in 2015 and 2018. The inclusion of more 15-year-olds in the assessed population often involves the inclusion of weaker students who would not have been enrolled or who would not have been at the appropriate grade level in earlier rounds of PISA. Once changes in coverage were accounted for, the average trend amongst the median and higher percentiles of 15-year-olds were not significant, although positive. It is therefore possible that the decline in mean performance in Costa Rica was due primarily to increased coverage of the 15-year-old population.

Snapshot of performance trends in CROATIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	477	467	493*
PISA 2009	476	460	486*
PISA 2012	485	471	491*
PISA 2015	487	464	475
PISA 2018	479	464	472
Average 3-year trend in mean performance	+1.4	-0.2	-5.3*
Short-term change in mean performance (2015 to 2018)	-7.9	+0.2	-3.0
Overall performance trajectory	flat	flat	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.5*	-1.8	-1.5*
Percentage-point change in low-achieving students (below Level 2)	-0.9	+1.3	+8.4*
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.9	+0.6	-2.9
Average trend amongst the lowest-achieving students (10th percentile)	+1.4	-0.9	-7.4*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In reading and mathematics, mean performance in Croatia remained stable, around a flat trend line, throughout the country's participation in PISA (2006-2018). In science, mean performance declined over this same period by about 5 score points on average per 3-year period. Performance declines in science were particularly pronounced amongst the country's lowest-achieving students. The proportion of students scoring below Level 2 in science increased by about 8 percentage points over that observed in PISA 2006.

Snapshot of performance trends in CYPRUS

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	m	m	m
PISA 2012	449*	440*	438
PISA 2015	443*	437*	433*
PISA 2018	424	451	439
Average 3-year trend in mean performance	-12.2*	+5.7*	+0.7
Short-term change in mean performance (2015 to 2018)	-18.5*	+13.6*	+6.4*
Overall performance trajectory	declining	improving	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-2.2*	+0.7	-0.3
Percentage-point change in low-achieving students (below Level 2)	+10.9*	-5.2*	+0.9
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-14.7*	+5.6*	+0.7
Average trend amongst the lowest-achieving students (10th percentile)	-0.7	+2.6	+3.1
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Cyprus participated in PISA for the third time in 2018. Mean reading performance declined over time, while mathematics performance improved between 2012 and 2018 and science performance returned close to the level observed in 2012.

Snapshot of performance trends in the CZECH REPUBLIC

Mean performance	Reading	Mathematics	Science
PISA 2000	492		
PISA 2003	489	516*	
PISA 2006	483	510	513*
PISA 2009	478*	493	500
PISA 2012	493	499	508*
PISA 2015	487	492	493
PISA 2018	490	499	497
Average 3-year trend in mean performance	+0.1	-3.7*	-4.0*
Short-term change in mean performance (2015 to 2018)	+3.0	+7.1	+4.0
Overall performance trajectory	U-shaped (more positive over more recent years)	negative, but flattening (less negative over more recent years)	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.1*	-0.2	-4.1*
Percentage-point change in low-achieving students (below Level 2)	-2.3	-0.6	+3.2
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.7	-5.9*	-4.9*
Average trend amongst the lowest-achieving students (10th percentile)	+0.2	-2.0	-3.2
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In the Czech Republic, mean performance in reading in 2018 was close to the level observed in all other PISA assessments since 2000, except PISA 2009. In mathematics, performance was below that observed in 2003, but above PISA 2015 performance. In science, performance was below that observed in 2006, but not significantly different from that observed in more recent years.

Snapshot of performance trends in DENMARK

Mean performance	Reading	Mathematics	Science
PISA 2000	497		
PISA 2003	492	514	
PISA 2006	494	513	496
PISA 2009	495	503	499
PISA 2012	496	500*	498
PISA 2015	500	511	502*
PISA 2018	501	509	493
Average 3-year trend in mean performance	+1.1	-0.9	-0.4
Short-term change in mean performance (2015 to 2018)	+1.3	-1.7	-9.3*
Overall performance trajectory	flat	U-shaped (more positive over more recent years)	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.7*	+1.7	-1.3
Percentage-point change in low-achieving students (below Level 2)	+0.8	-2.3	+0.2
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.7	-3.4*	-1.2
Average trend amongst the lowest-achieving students (10th percentile)	+2.2	+1.0	+0.2
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in reading remained stable, around a flat trend line, throughout Denmark's participation in PISA. In mathematics and science too, no overall direction of the trend could be detected; however, in mathematics, a declining trend up to 2012 was followed by a (partial) recovery over the 2012-2018 period, while in science, performance in 2018 was about 9 score points lower, on average, than in 2015. The overall trend in mathematics performance was negative amongst the highest-achieving students (at the 90th percentile).

Snapshot of performance trends in the DOMINICAN REPUBLIC

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	m	m	m
PISA 2012	m	m	m
PISA 2015	358*	328	332
PISA 2018	342	325	336
Average 3-year trend in mean performance	-16.1*	-2.6	+4.0
Short-term change in mean performance (2015 to 2018)	-16.1*	-2.6	+4.0
Overall performance trajectory	declining	stable	stable
Proficiency levels	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.0	+0.0	-0.0
Percentage-point change in low-achieving students (below Level 2)	+6.9*	+0.0	-0.9
Variation in performance	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-17.5*	-0.4	+2.0
Average trend amongst the lowest-achieving students (10th percentile)	-9.4	-6.4	+6.2
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

The Dominican Republic participated in PISA for the second time since 2015. While mathematics and science performance was similar to that observed in 2015, reading performance lay 16 score points below that observed in 2015.

Snapshot of performance trends in ESTONIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	501*	515	531
PISA 2009	501*	512*	528
PISA 2012	516	521	541*
PISA 2015	519	520	534
PISA 2018	523	523	530
Average 3-year trend in mean performance	+6.3*	+2.5*	+0.4
Short-term change in mean performance (2015 to 2018)	+3.9	+3.9	-4.1
Overall performance trajectory	steadily positive	steadily positive	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+7.8*	+0.9	+0.7
Percentage-point change in low-achieving students (below Level 2)	-2.3	-0.3	+1.1
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+10.0*	+2.6*	+2.1
Average trend amongst the lowest-achieving students (10th percentile)	+3.7*	+2.2	-1.3
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean reading and mathematics performance in Estonia improved steadily since the country first participated in PISA in 2006. Over this same period (2006-2018), performance in science remained mostly stable (and high). The improvement in reading performance was particularly marked at the top of the performance distribution: the 90th percentile moved up on the PISA scale by about 10 points every 3 years and, between 2009 and 2018, the proportion of student scoring at Level 5 or 6 (top performers) increased by almost 8 percentage points.

Snapshot of performance trends in FINLAND

Mean performance	Reading	Mathematics	Science
PISA 2000	546*		
PISA 2003	543*	544*	
PISA 2006	547*	548*	563*
PISA 2009	536*	541*	554*
PISA 2012	524	519*	545*
PISA 2015	526	511	531*
PISA 2018	520	507	522
Average 3-year trend in mean performance	-4.9*	-9.1*	-10.7*
Short-term change in mean performance (2015 to 2018)	-6.3	-3.8	-8.8*
Overall performance trajectory	steadily negative	increasingly negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.3	-4.1*	-8.6*
Percentage-point change in low-achieving students (below Level 2)	+5.4*	+2.7*	+8.8*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-1.5	-9.3*	-7.2*
Average trend amongst the lowest-achieving students (10th percentile)	-8.6*	-9.7*	-15.5*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean reading, mathematics and science performance continued to decline in Finland. In all three subjects the decline began after 2006. Although PISA 2018 results were significantly lower than PISA 2015 results only in science, there was no sign of a flattening or reversing trend in any subject. In mathematics, declines were similarly rapid at all levels of the performance distribution; in reading and science, in contrast, the declining trend was particularly noticeable amongst the lowest-achieving students. The proportion of top-performing students in mathematics shrank by 4 percentage points between 2012 and 2018, while the proportion of top-performing students in science decreased by 9 percentage points between 2006 and 2018. Meanwhile, the proportion of low-achieving students in reading grew by 5 percentage points between 2009 and 2018; the proportion of low-achieving students in mathematics grew by 3 percentage points between 2012 and 2018; and the share of low performers in science increased by 9 percentage points between 2006 and 2018.

Snapshot of performance trends in FRANCE

Mean performance	Reading	Mathematics	Science
PISA 2000	505*		
PISA 2003	496	511*	
PISA 2006	488	496	495
PISA 2009	496	497	498
PISA 2012	505*	495	499
PISA 2015	499	493	495
PISA 2018	493	495	493
Average 3-year trend in mean performance	-0.4	-2.5*	-0.8
Short-term change in mean performance (2015 to 2018)	-6.7	+2.5	-2.0
Overall performance trajectory	flat	negative, but flattening (less negative over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.4	-1.9	-1.5
Percentage-point change in low-achieving students (below Level 2)	+1.2	-1.1	-0.7
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+3.0*	-2.8*	-1.7
Average trend amongst the lowest-achieving students (10th percentile)	-4.0*	-3.1*	+0.7
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean science performance in France remained stable over the 2006-2018 period; similarly, no overall direction of change can be determined for mean reading performance over the 2000-2018 period. Mathematics performance declined between 2003 and 2018, but most of that decline was observed in earlier assessments; the recent trend is flat in mathematics too.

In reading, the apparent stability hides distinct trends amongst students at different levels in the performance distribution. Amongst the lowest-achieving students, performance tended to decline (by 4 score points, on average, per 3-year period); whereas amongst the highest-achieving students, performance tended to improve (by 3 score points, on average, per 3-year-period). No such widening of performance gaps was observed in mathematics (where a similar decline was observed amongst the highest-and lowest-achieving students, on average) and science.

Snapshot of performance trends in GEORGIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	374	379*	373
PISA 2012	m	m	m
PISA 2015	401*	404	411*
PISA 2018	380	398	383
Average 3-year trend in mean performance	+3.5	+7.6*	+5.6*
Short-term change in mean performance (2015 to 2018)	-21.5*	-6.2	-28.5*
Overall performance trajectory	stable	improving	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.1	m	-0.1
Percentage-point change in low-achieving students (below Level 2)	+2.4	m	-1.2
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.5	+11.2*	+3.4
Average trend amongst the lowest-achieving students (10th percentile)	+11.6*	+5.9*	+10.6*
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

PISA 2018 results in Georgia were significantly below those observed in 2015 in reading and science, reversing most of the gains observed between 2010 and 2015. Only mathematics results in PISA 2018 remained significantly above the level observed in 2010.

Snapshot of performance trends in GERMANY

Mean performance	Reading	Mathematics	Science
PISA 2000	484*		
PISA 2003	491	503	
PISA 2006	495	504	516*
PISA 2009	497	513*	520*
PISA 2012	508	514*	524*
PISA 2015	509	506	509
PISA 2018	498	500	503
Average 3-year trend in mean performance	+3.3*	-0.1	-3.6*
Short-term change in mean performance (2015 to 2018)	-10.8	-5.9	-6.2
Overall performance trajectory	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)	negative, and more so over more recent years
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.7*	-4.1*	-1.8
Percentage-point change in low-achieving students (below Level 2)	+2.2	+3.4*	+4.2*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+1.9	-2.8*	-2.6
Average trend amongst the lowest-achieving students (10th percentile)	+5.8*	+2.8	-4.2*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Germany, mean reading and mathematics performance in 2018 returned close to levels that were last observed in 2006 or 2009, reversing most of the gains observed over the early period (up to 2012); in science, mean performance was below 2006 levels. PISA 2018 results lay significantly below PISA 2012 results in mathematics.

The recent trajectory of mean reading performance could be partly related to the changing composition of the student population. It could be estimated that, if the student population in 2015 had had the same demographic profile as the population in 2018, the average score in reading would have been 505 points (Table I.B1.40), or about 5 score points below the average observed score (Table I.B1.10). However, demographic changes account only for a small part of the larger negative trends observed in mathematics and science since 2012.

In mathematics, while there was no overall trend in mean performance over the full 2003-2018 period, the trend was negative amongst the highest-achieving students (those at the 90th percentile).

Over the most recent period, performance trends in Germany differed by gender. Between 2015 and 2018, girls' performance in mathematics and science remained stable, while mean score amongst boys declined by 11 points in mathematics and by 12 points in science (Tables II.B1.7.36 and II.B1.7.42 in *PISA 2018 Results [Volume II]: Where All Students Can Succeed*).

Snapshot of performance trends in GREECE

Mean performance	Reading	Mathematics	Science
PISA 2000	474*		
PISA 2003	472	445	
PISA 2006	460	459	473*
PISA 2009	483*	466*	470*
PISA 2012	477*	453	467*
PISA 2015	467	454	455
PISA 2018	457	451	452
Average 3-year trend in mean performance	-1.5	+0.1	-5.9*
Short-term change in mean performance (2015 to 2018)	-9.6	-2.3	-3.2
Overall performance trajectory	hump-shaped (more negative over more recent years)	hump-shaped (more negative over more recent years)	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-2.0*	-0.2	-2.1*
Percentage-point change in low-achieving students (below Level 2)	+9.2*	+0.1	+7.7*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-1.5	-0.8	-6.4*
Average trend amongst the lowest-achieving students (10th percentile)	-0.8	+0.5	-5.3*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean science performance in Greece declined steadily since 2006, by an average of 5.9 score points per 3-year period, even though changes from one round to the next were not always statistically significant. Performance in mathematics can be described as hump-shaped, mainly due to a spike in performance in PISA 2009; performance in other years was stable. Similarly, mean reading performance can be described as hump-shaped, with a steady decline in performance since its peak in 2009. Greece performed below the OECD average in all subjects in every year it participated in PISA.

The decline in science performance over the 2006-2018 period was observed across the performance distribution. Performance amongst the highest-achieving students declined by 6.4 percentage points and that amongst the lowest-achieving students fell by 5.3 percentage points per 3-year period.

Snapshot of performance trends in HONG KONG (CHINA)

Mean performance	Reading	Mathematics	Science
PISA 2000	525		
PISA 2003	510	550	
PISA 2006	536	547	542*
PISA 2009	533	555	549*
PISA 2012	545*	561	555*
PISA 2015	527	548	523
PISA 2018	524	551	517
Average 3-year trend in mean performance	+1.6	+0.4	-7.7*
Short-term change in mean performance (2015 to 2018)	-2.4	+3.2	-6.6
Overall performance trajectory	hump-shaped (more negative over more recent years)	flat	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.4	-4.7*	-8.1*
Percentage-point change in low-achieving students (below Level 2)	+4.3*	+0.7	+2.8*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.8*	-1.0	-9.6*
Average trend amongst the lowest-achieving students (10th percentile)	-1.5	+1.6	-5.4*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Hong Kong (China), mean reading, mathematics and science performance in 2018 was close to the level observed in 2015. When considering a longer period, reading performance in 2015-2018 was below PISA 2012 levels, but not significantly different from 2009 or 2002, the previous years in which reading was the major focus of the assessment. Science performance was below the level observed over the 2006-2012 period, while mathematics performance appeared stable, fluctuating around a flat trend over the 2003-2018 period.

The apparent stability in reading performance between 2002, 2009 and 2018, however, hides widening performance gaps between the highest- and the lowest-achieving students. No similar widening of performance gaps was observed in either mathematics or science.

In reading, the proportion of students scoring below Level 2 (low-achieving students) increased by 4 percentage points between 2009 and 2018. In science, the proportion of top-performing students decreased by 8 percentage points between 2006 and 2018.

Snapshot of performance trends in HUNGARY

Mean performance	Reading	Mathematics	Science
PISA 2000	480		
PISA 2003	482	490	
PISA 2006	482	491*	504*
PISA 2009	494*	490	503*
PISA 2012	488*	477	494*
PISA 2015	470	477	477
PISA 2018	476	481	481
Average 3-year trend in mean performance	-1.1	-2.8*	-7.1*
Short-term change in mean performance (2015 to 2018)	+6.5	+4.3	+4.2
Overall performance trajectory	hump-shaped (more negative over more recent years)	steadily negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.4	-1.3	-2.2*
Percentage-point change in low-achieving students (below Level 2)	+7.7*	-2.4	+9.1*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.3	-3.0*	-3.6*
Average trend amongst the lowest-achieving students (10th percentile)	-2.4	-3.7*	-10.6*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Hungary's average performance in reading in 2018 was close to its level in 2000, when the country first participated in PISA; but as is also observed in science and to a lesser extent in mathematics, the more recent trend, after 2009, was negative. In particular, the proportion of low-achieving students (students scoring below Level 2) increased by about 8 percentage points in reading (2009-2018) and by about 9 percentage points in science (2006-2018).

Snapshot of performance trends in ICELAND

Mean performance	Reading	Mathematics	Science
PISA 2000	507*		
PISA 2003	492*	515*	
PISA 2006	484	506*	491*
PISA 2009	500*	507*	496*
PISA 2012	483	493	478
PISA 2015	482	488*	473
PISA 2018	474	495	475
Average 3-year trend in mean performance	-4.4*	-4.7*	-5.4*
Short-term change in mean performance (2015 to 2018)	-7.6	+7.2*	+1.8
Overall performance trajectory	steadily negative	negative, but flattening (less negative over more recent years)	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-1.4	-0.8	-2.5*
Percentage-point change in low-achieving students (below Level 2)	+9.5*	-0.8	+4.4*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-1.7	-4.1*	-6.1*
Average trend amongst the lowest-achieving students (10th percentile)	-6.5*	-5.6*	-3.8*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in all three subjects declined over Iceland's participation in PISA by about 5 score points per 3-year period, on average. While, in mathematics, mean performance in 2018 was higher than that observed in 2015, reversing some earlier losses, this was not observed in reading or in science. Performance in reading declined amongst the country's lowest-achieving students (at the 10th percentile), while no decline was observed amongst the highest-achieving students (at the 90th percentile). The proportion of students who scored below Level 2 in reading increased by 9.5 percentage points between 2009 and 2018.

Snapshot of performance trends in INDONESIA

Mean performance	Reading	Mathematics	Science
PISA 2000	371		
PISA 2003	382	360*	
PISA 2006	393*	391	393
PISA 2009	402*	371	383*
PISA 2012	396*	375	382*
PISA 2015	397*	386	403
PISA 2018	371	379	396
Average 3-year trend in mean performance	+1.2	+2.2	+2.5
Short-term change in mean performance (2015 to 2018)	-26.3*	-7.4	-7.0
Overall performance trajectory	hump-shaped (more negative over more recent years)	hump-shaped (more negative over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.0	+0.2	+0.0
Percentage-point change in low-achieving students (below Level 2)	+16.5*	-3.8	-1.6
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.1	+1.5	+1.9
Average trend amongst the lowest-achieving students (10th percentile)	+1.2	+2.7	+3.0*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Indonesia has participated in PISA since 2001. Since that time, performance in science has fluctuated but remained flat overall, while performance in both reading and mathematics has been hump-shaped. Reading performance in 2018 fell back to its 2001 level after a peak in 2009, while mathematics performance fluctuated more in the early years of PISA but remained relatively stable since 2009.

However, these results must be seen in the context of the vast strides that Indonesia has made in increasing enrolment. In 2003, the PISA sample covered only 46% of 15-year-olds in Indonesia; in 2018, 85% of 15-year-olds were covered. It is often the case that the strongest students remain in education, and that students who were not in education and were brought into the school system are weaker than those who were already included. If there had been no improvement in the education system, the inclusion of more students would be expected to lower mean performance and the performance distribution. In that light, in maintaining education standards over its participation in PISA, Indonesia has been able to raise the quality of its education system.

Trends adjusted for enrolment show this more clearly. On the assumption that the 15-year-olds who were excluded from the PISA sample would have performed below the 75th percentile of all 15-year-olds if they had sat the assessment, the mathematics and science performance of the highest-achieving 25% amongst all 15-year-olds in Indonesia would have improved by 11 points every three years since 2003 (Tables I.B1.35 and I.B1.36).

Snapshot of performance trends in IRELAND

Mean performance	Reading	Mathematics	Science
PISA 2000	527		
PISA 2003	515	503	
PISA 2006	517	501	508*
PISA 2009	496*	487*	508*
PISA 2012	523	501	522*
PISA 2015	521	504	503
PISA 2018	518	500	496
Average 3-year trend in mean performance	-0.3	+0.1	-3.0*
Short-term change in mean performance (2015 to 2018)	-2.7	-4.1	-6.5
Overall performance trajectory	U-shaped (more positive over more recent years)	U-shaped (more positive over more recent years)	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+5.1*	-2.4*	-3.6*
Percentage-point change in low-achieving students (below Level 2)	-5.4*	-1.2	+1.5
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-0.2	-1.8	-5.0*
Average trend amongst the lowest-achieving students (10th percentile)	+0.6	+1.3	-0.7
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

PISA 2018 results in Ireland were close to their historic average in reading and mathematics, with no significant overall direction of change. While the trajectory of reading and mathematics performance can be described as U-shaped, this is entirely the result of PISA 2009 results, which were significantly below the historic average. Mean performance in all other years was close to that observed in PISA 2018.

In science, the overall trend was negative; in particular, the more recent trend (since 2012) and the trend amongst the highest-performing students was markedly negative. Between 2006 and 2018, the proportion of students who scored at Level 5 or 6 on the PISA scale (top-performing students) decreased by 3.6 percentage points, and the 90th percentile of the performance distribution moved down on the PISA scale by about 5 score points per 3-year period.

Snapshot of performance trends in ISRAEL

Mean performance	Reading	Mathematics	Science
PISA 2000	452		
PISA 2003	m	m	
PISA 2006	439*	442*	454
PISA 2009	474	447*	455
PISA 2012	486*	466	470
PISA 2015	479	470	467
PISA 2018	470	463	462
Average 3-year trend in mean performance	+6.1*	+6.4*	+2.8
Short-term change in mean performance (2015 to 2018)	-8.5	-6.6	-4.4
Overall performance trajectory	positive, but flattening (less positive over more recent years)	positive, but flattening (less positive over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.0*	-0.6	+0.6
Percentage-point change in low-achieving students (below Level 2)	+4.5*	+0.6	-3.0
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+8.7*	+5.8*	+2.9
Average trend amongst the lowest-achieving students (10th percentile)	+2.6	+4.4*	+2.0
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Performance improved in reading (since 2001) and mathematics (since 2006) in Israel, although most of that improvement happened in the early period (up to 2012). Since 2012, no significant changes were observed in mathematics performance, while reading performance declined somewhat. Performance in science remained stable throughout the 2006-2018 period.

Over the 2001-2018 period, improvements in reading performance were particularly marked amongst the highest-achieving students. The 90th percentile, i.e. the level above which only 10% of all students scored, increased by 8.7 score points per 3-year period, significantly faster than the 10th percentile. As a result, performance gaps in reading widened.

Snapshot of performance trends in ITALY

Mean performance	Reading	Mathematics	Science
PISA 2000	487*		
PISA 2003	476	466*	
PISA 2006	469	462*	475
PISA 2009	486*	483	489*
PISA 2012	490*	485	494*
PISA 2015	485	490	481*
PISA 2018	476	487	468
Average 3-year trend in mean performance	+0.2	+5.4*	-2.3
Short-term change in mean performance (2015 to 2018)	-8.5	-3.1	-12.5*
Overall performance trajectory	flat	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.5	-0.4	-1.9*
Percentage-point change in low-achieving students (below Level 2)	+2.2	-0.8	+0.6
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.4	+4.6*	-4.3*
Average trend amongst the lowest-achieving students (10th percentile)	+0.1	+5.2*	-0.9
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Italy, mean reading performance in 2018 was below the level observed in PISA 2000 and PISA 2009 (the two prior assessments with reading as the main focus), but close to the level observed in most remaining assessments, and no clear direction of change could be determined. Mean science performance in 2018 was significantly below the level observed over the 2009-2015 period, and returned to a level last observed in 2006. Mean mathematics performance in Italy improved in the early cycles of PISA, then remained stable after 2009.

Over the 2006-2018 period, science performance declined most markedly amongst the highest-achieving students. The 90th percentile of performance in science, i.e. the level above which only 10% of all students scored, declined by 4.3 score points per 3-year period, significantly faster than the 10th percentile. As a result, performance gaps in science narrowed, and the proportion of students who scored at Level 5 or 6 in science (top-performing students) shrank by 1.9 percentage points.

Snapshot of performance trends in JAPAN

Mean performance	Reading	Mathematics	Science
PISA 2000	522*		
PISA 2003	498	534	
PISA 2006	498	523	531
PISA 2009	520*	529	539
PISA 2012	538*	536	547*
PISA 2015	516*	532	538*
PISA 2018	504	527	529
Average 3-year trend in mean performance	+0.8	-0.0	-0.6
Short-term change in mean performance (2015 to 2018)	-12.1*	-5.5	-9.3*
Overall performance trajectory	flat	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-3.2*	-5.3*	-2.0
Percentage-point change in low-achieving students (below Level 2)	+3.2	+0.4	-1.2
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+1.8	-2.7	-2.2
Average trend amongst the lowest-achieving students (10th percentile)	+0.9	+2.9	+2.3
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean mathematics performance in Japan remained stable over the 2003-2018 period, with no significant improvement or deterioration over any sub-period. However, this apparent stability hides distinct trends amongst students at different levels in the performance distribution. Amongst the highest-achieving students in particular, performance tended to decline (by 2.7 score points, on average, per 3-year period; although this trend is not significantly different from 0, it is significantly different from the trend observed amongst the lowest-achieving students).

While no overall direction of change can be determined for reading and science trends in Japan, mean performance in these subjects has been characterised by significant instability. Results appeared more stable when considering only years in which each subject was assessed fully (2000, 2009 and 2018 for reading; 2006 and 2015 for science), perhaps indicating that some of this instability is related to the change in subject coverage in the "off" years (such changes were particularly marked in PISA cycles prior to 2015). Even so, in reading, the more recent trend (since 2009 or 2015) was clearly negative. In science too, mean performance in 2018 was below Japan's performance in PISA 2012 and 2015.

Similar to mathematics, trends amongst the highest-performing students in science tend to be more negative than amongst the lowest-performing students. This narrowing gap in performance is not observed in reading.

Snapshot of performance trends in JORDAN

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	401*	384*	422
PISA 2009	405*	387*	415*
PISA 2012	399*	386*	409*
PISA 2015	408	380*	409*
PISA 2018	419	400	429
Average 3-year trend in mean performance	+4.0*	+2.5	+0.8
Short-term change in mean performance (2015 to 2018)	+11.0	+19.5*	+20.6*
Overall performance trajectory	increasingly positive	U-shaped (more positive over more recent years)	U-shaped (more positive over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.0	+0.1	+0.0
Percentage-point change in low-achieving students (below Level 2)	-6.8	-9.2*	-4.0
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.6	+3.6*	-0.1
Average trend amongst the lowest-achieving students (10th percentile)	+4.9*	+1.6	+1.1
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Jordan, mean performance improved in all three subjects over the more recent assessments (i.e. since 2012 or 2015), after initially flat or even declining (mathematics) trends between 2006 and 2012 (the overall trajectory of performance, since 2006, is significantly positive only in reading).

However, these positive trends since 2012 were observed during a period in which enrolment rates for 15-year-olds in grade 7 and above did not keep pace with increases in the resident population of 15-year-olds. While the population of 15-year-olds enrolled in grade 7 and above, and represented by PISA samples, remained close to the level observed in 2012, the overall population of 15-year-olds increased by more than 25% over the same period, largely as a result of a massive influx of refugees from neighbouring countries. Refugee children may be enrolled outside of Jordan’s formal education system.

Snapshot of performance trends in KAZAKHSTAN

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	390	405*	400
PISA 2012	393	432	425*
PISA 2015	m	m	m
PISA 2018	387	423	397
Average 3-year trend in mean performance	-1.4	+4.7*	-2.9
Short-term change in mean performance (2015 to 2018)	m	m	m
Overall performance trajectory	stable	improving	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.0	+1.0*	+0.1
Percentage-point change in low-achieving students (below Level 2)	+5.5	+3.9	+4.9
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-6.2*	+6.7*	-6.6*
Average trend amongst the lowest-achieving students (10th percentile)	+5.0*	+1.3	+2.2
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Kazakhstan's mean performance in reading and science in 2018 was close to the level observed in 2009, when the country first participated in PISA. In contrast, in mathematics, mean performance showed significant improvements from the 2009 level. Mathematics performance improved, particularly amongst the highest-performing students; and the share of students who scored at Level 5 or 6 in mathematics increased by 1 percentage point between 2012 and 2018. At the same time, performance in reading and science declined amongst the highest-performing students.

PISA 2015 results for Kazakhstan cannot be compared to results from previous years or to those from 2018 due to the potential of bias introduced by incomplete student-response data. PISA 2018 results fully met the technical standards.

Snapshot of performance trends in KOREA

Mean performance	Reading	Mathematics	Science
PISA 2000	525		
PISA 2003	534*	542*	
PISA 2006	556*	547*	522
PISA 2009	539*	546*	538*
PISA 2012	536*	554*	538*
PISA 2015	517	524	516
PISA 2018	514	526	519
Average 3-year trend in mean performance	-3.1*	-4.1*	-2.9*
Short-term change in mean performance (2015 to 2018)	-3.4	+1.8	+3.2
Overall performance trajectory	increasingly negative	increasingly negative	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.2	-9.5*	+1.5
Percentage-point change in low-achieving students (below Level 2)	+9.3*	+5.9*	+2.9
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.6	-1.9	+1.0
Average trend amongst the lowest-achieving students (10th percentile)	-9.5*	-7.3*	-7.6*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Korea, mean reading, mathematics and science performance in 2018 was close to the level observed in 2015, and below the level observed in 2009 and 2012. In reading and science, this recent decline in performance reversed earlier gains.

Across all three subjects, a significant widening of performance differences could be observed. While no decline was observed amongst the highest-achieving students (the level above which only 10% of students scored remained stable), the lowest-achieving students lost significant ground in all subjects over the period. The 10th percentile of the distribution, representing the level above which 90% of students scored, declined by more than 7 points, on average, per 3-year period, or more than 20 points per decade.

Snapshot of performance trends in KOSOVO

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	m	m	m
PISA 2012	m	m	m
PISA 2015	347	362	378*
PISA 2018	353	366	365
Average 3-year trend in mean performance	+5.9	+4.4	-13.6*
Short-term change in mean performance (2015 to 2018)	+5.9	+4.4	-13.6*
Overall performance trajectory	stable	stable	declining
Proficiency levels	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.0	+0.1	-0.0
Percentage-point change in low-achieving students (below Level 2)	+1.9	-1.1	+8.8*
Variation in performance	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-4.6	+5.3	-24.1*
Average trend amongst the lowest-achieving students (10th percentile)	+22.0*	+3.1	-3.5
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Kosovo participated in PISA for the second time in 2018. Mean performance was similar in reading and mathematics, but was 14 points lower in science, than in 2015. In science, the proportion of students scoring below Level 2 increased by 9 percentage points over the period.

Snapshot of performance trends in LATVIA

Mean performance	Reading	Mathematics	Science
PISA 2000	458*		
PISA 2003	491	483*	
PISA 2006	479	486*	490
PISA 2009	484	482*	494
PISA 2012	489*	491	502*
PISA 2015	488*	482*	490
PISA 2018	479	496	487
Average 3-year trend in mean performance	+2.3	+1.7	-0.8
Short-term change in mean performance (2015 to 2018)	-9.1*	+13.8*	-3.0
Overall performance trajectory	hump-shaped (more negative over more recent years)	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.9*	+0.5	-0.4
Percentage-point change in low-achieving students (below Level 2)	+4.9*	-2.6	+1.1
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.3	+0.1	+0.1
Average trend amongst the lowest-achieving students (10th percentile)	+4.7*	+3.5*	-1.4
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

When taking into account results from all years, no significant improving or declining trend could be determined, in any subject, in Latvia. In 2018, mean reading performance in Latvia was above the level observed when the country first participated in PISA in 2000, but below the level observed in 2015. Mean mathematics performance was significantly higher in PISA 2018 than in PISA 2015, but when considering the entire 2003-2018 period, mathematics performance appeared to oscillate around a stable mean, with no clear direction of change. Science performance in PISA 2018 was close to that observed in all previous assessments, except in 2012.

A more consistently positive trend was observed amongst the lowest-achieving students in reading and mathematics, narrowing the gap between those and higher-achieving students to some extent.

Snapshot of performance trends in LEBANON

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	m	m	m
PISA 2012	m	m	m
PISA 2015	347	396	386
PISA 2018	353	393	384
Average 3-year trend in mean performance	+6.8	-2.8	-2.8
Short-term change in mean performance (2015 to 2018)	+6.8	-2.8	-2.8
Overall performance trajectory	stable	stable	stable
Proficiency levels	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.1	+0.1	+0.1
Percentage-point change in low-achieving students (below Level 2)	-2.6	-0.5	-0.4
Variation in performance	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.0	+1.4	+2.6
Average trend amongst the lowest-achieving students (10th percentile)	+8.4	-11.9	-10.6
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

PISA 2018 results for Lebanon, in all three subjects, were close to those observed in 2015, when the country first participated in PISA. This stability of results is remarkable because the proportion of 15-year-olds who were eligible to participate in the PISA assessment increased by about 25% since 2015 (Table I.A2.2).

Snapshot of performance trends in LITHUANIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	470	486	488
PISA 2009	468	477	491
PISA 2012	477	479	496*
PISA 2015	472	478	475
PISA 2018	476	481	482
Average 3-year trend in mean performance	+1.6	-0.7	-2.8*
Short-term change in mean performance (2015 to 2018)	+3.5	+2.8	+6.7
Overall performance trajectory	flat	U-shaped (more positive over more recent years)	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.1*	+0.4	-0.5
Percentage-point change in low-achieving students (below Level 2)	+0.0	-0.4	+1.8
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.7	-0.8	-1.5
Average trend amongst the lowest-achieving students (10th percentile)	+1.0	-0.9	-3.7*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Lithuania, mean reading and mathematics performance in 2018 were close to the levels observed in every previous assessments since 2006, when the country first participated in PISA, and no clear direction of change could be determined. Mean science performance in 2018 was significantly above the level observed in 2015, but below the PISA 2012 mean; overall, science results appeared to fluctuate somewhat more than reading or mathematics results, around a declining trend. Despite overall stable results in reading, the proportion of top-performing students increased by 2.1 percentage points between 2009 and 2018.

Snapshot of performance trends in LUXEMBOURG

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	479	493*	
PISA 2006	479	490	486*
PISA 2009	472	489	484
PISA 2012	488*	490	491*
PISA 2015	481*	486	483*
PISA 2018	470	483	477
Average 3-year trend in mean performance	-0.7	-1.7	-1.9
Short-term change in mean performance (2015 to 2018)	-11.5*	-2.3	-6.0*
Overall performance trajectory	hump-shaped (more negative over more recent years)	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.9*	-0.4	-0.4
Percentage-point change in low-achieving students (below Level 2)	+3.3*	+2.9	+4.7*
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+3.5*	-0.3	-0.6
Average trend amongst the lowest-achieving students (10th percentile)	-2.9*	-3.1*	-1.4
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in Luxembourg remained largely stable in mathematics since 2003, although mean performance was 10 score points higher in 2003 than in 2018. Mean performance in reading and science, in contrast, was lower in 2018 than in the most recent previous assessments (2012 and 2015): performance declined by 11 score points in reading and by 6 score points in science between 2015 and 2018.

Between 2009 and 2018, the proportion of 15-year-old students who reported an immigrant background in Luxembourg increased by 15 percentage points, the largest increase amongst OECD countries (Table II.B1.9.9 in *PISA 2018 Results (Volume II): Where All Students Can Succeed*). While immigrant students continued to score more than 30 points below non-immigrant students in reading, performance amongst immigrant students improved significantly between 2009 and 2018 (Table II.B1.9.10). Nevertheless, the change in the proportion of immigrant and non-immigrant students alone could account for about five points (15% of 30 points) of the 18-point decline in mean reading scores over the 2012-2018 period (see also Table I.B1.40).

The gap in performance between the highest- and lowest-achieving students in Luxembourg increased in both reading and mathematics since 2003. Higher shares of immigrant students likely contributed to this widening trend. It can be estimated that, if the student population in 2009 had had the same demographic characteristics as the student population in 2018, no widening of the gap in reading performance would have been observed between 2009 and 2018 (Table I.B1.46).

Snapshot of performance trends in MACAO (CHINA)

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	498*	527*	
PISA 2006	492*	525*	511*
PISA 2009	487*	525*	511*
PISA 2012	509*	538*	521*
PISA 2015	509*	544*	529*
PISA 2018	525	558	544
Average 3-year trend in mean performance	+6.0*	+6.2*	+8.3*
Short-term change in mean performance (2015 to 2018)	+16.4*	+13.9*	+15.0*
Overall performance trajectory	increasingly positive	increasingly positive	increasingly positive
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+10.9*	+3.3	+8.3*
Percentage-point change in low-achieving students (below Level 2)	-4.1*	-5.8*	-4.3*
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+11.2*	+4.5*	+9.7*
Average trend amongst the lowest-achieving students (10th percentile)	-0.1	+7.4*	+6.0*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	narrowing gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Macao (China) showed increasingly positive trends in mean performance in all three subjects over its participation in PISA. Performance in reading and mathematics improved by about 6 score points per 3-year period since 2003; performance in science improved by 8.3 score points per 3-year period since 2006. Improvements were even larger between 2015 and 2018, exceeding 13 score points in all three subjects.

Improvements in reading and science were particularly strong amongst the highest-achieving students. In contrast, in mathematics, improvements were more rapid amongst the lowest-achieving students. The proportion of students performing below Level 2 shrank in all three subjects (reading, mathematics and science), while the proportion of students performing at Level 5 or 6 increased in reading (by about 11 percentage points) and science (by about 8 percentage points). These are amongst the most rapid improvements observed amongst all PISA-participating countries and economies.

Snapshot of performance trends in MALAYSIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	414	404*	422*
PISA 2012	398*	421*	420*
PISA 2015	m	m	m
PISA 2018	415	440	438
Average 3-year trend in mean performance	+2.2	+12.7*	+6.6*
Short-term change in mean performance (2015 to 2018)	m	m	m
Overall performance trajectory	stable	improving	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.4	+1.1*	+0.4
Percentage-point change in low-achieving students (below Level 2)	+1.9	-10.3*	-6.4
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+5.6*	+16.8*	+7.5*
Average trend amongst the lowest-achieving students (10th percentile)	+1.1	+8.7*	+6.5*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Malaysia, mean mathematics and science performance in 2018 lay above the performance observed in 2009, when the country first participated in PISA, and in 2012. In reading, performance in 2018 was close to that observed in 2009 but better than that observed in 2012. Improvements were observed, in general, amongst both high- and low-achieving students; but improvements in mathematics were particularly strong amongst the country's highest-achieving students: at the 90th percentile, performance improved by about 17 score points per 3-year period.

PISA 2015 results for Malaysia cannot be compared to results from previous years or to those from 2018 due to the potential of bias introduced by low response rates in the original PISA sample. PISA 2018 results fully met the technical standards.

Snapshot of performance trends in MALTA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	442	463*	461
PISA 2012	m	m	m
PISA 2015	447	479*	465*
PISA 2018	448	472	457
Average 3-year trend in mean performance	+2.3	+3.9*	-1.3
Short-term change in mean performance (2015 to 2018)	+1.6	-6.9*	-8.2*
Overall performance trajectory	stable	improving	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.9	m	-1.6*
Percentage-point change in low-achieving students (below Level 2)	-0.4	m	+1.0
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.2	+2.4	-4.3*
Average trend amongst the lowest-achieving students (10th percentile)	+5.4	+3.1	+2.6
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Malta, mean performance in reading and science in PISA 2018 was close to that observed in 2010, when the country first participated in PISA. In mathematics, mean performance lay above the performance observed in 2010.

Snapshot of performance trends in MEXICO

Mean performance	Reading	Mathematics	Science
PISA 2000	422		
PISA 2003	400*	385*	
PISA 2006	410	406	410
PISA 2009	425	419*	416
PISA 2012	424	413	415
PISA 2015	423	408	416
PISA 2018	420	409	419
Average 3-year trend in mean performance	+2.0	+3.4*	+1.9
Short-term change in mean performance (2015 to 2018)	-2.8	+0.8	+3.5
Overall performance trajectory	flat	positive, but flattening (less positive over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.4	-0.1	-0.0
Percentage-point change in low-achieving students (below Level 2)	+4.6	+1.5	-4.1
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-0.4	+0.7	-0.2
Average trend amongst the lowest-achieving students (10th percentile)	+4.9*	+6.0*	+4.5*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in reading, mathematics and science in Mexico remained stable, around a flat trend line, throughout most of the country's participation in PISA. Only PISA 2003 performance (in reading and mathematics) was significantly below its 2018 level; in all other years, and across all subjects, mean performance was not significantly different from PISA 2018. However, this overall stability hides more positive trends amongst the lowest-achieving students. The score reached by at least 90% of students in Mexico (10th percentile) increased, on average, by about 5 score points per 3-year period in each of the three subjects (reading, mathematics and science). As a result of improvements amongst low-achieving students in mathematics and science, the gaps in performance between the highest- and lowest-achieving students in these two subjects shrank over time.

These performance trends were observed over a period of rapid expansion of secondary education. Between 2003 and 2018, Mexico added more than 400 000 students to the total population of 15-year-olds eligible to participate in PISA; the proportion of 15-year-olds who were covered by PISA samples increased from about 50% in 2003 to 66% in 2018. It is likely that this expansion in education opportunities dampened a more positive underlying trend in student performance. Indeed, a simulation that assumes that the highest-scoring 25% of 15-year-olds were eligible to take the test in any given year shows a positive trend amongst this population in mathematics (since 2003) and science (since 2006) (Figure I.9.5).

Snapshot of performance trends in the REPUBLIC OF MOLDOVA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	388*	397*	413*
PISA 2012	m	m	m
PISA 2015	416	420	428
PISA 2018	424	421	428
Average 3-year trend in mean performance	+13.7*	+9.2*	+6.1*
Short-term change in mean performance (2015 to 2018)	+7.8	+0.9	+0.5
Overall performance trajectory	improving	improving	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.9*	m	+0.6*
Percentage-point change in low-achieving students (below Level 2)	-14.2*	m	-4.7
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+17.0*	+13.6*	+8.6*
Average trend amongst the lowest-achieving students (10th percentile)	+11.1*	+5.0*	+5.9*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean reading, mathematics and science performance in the Republic of Moldova improved since the country first participated in PISA in 2010. In reading and science, improvements amongst the highest- and lowest-achieving students were similar, and close to the average improvements. In mathematics too, students at all levels improved their performance, but the highest-achieving students improved more than the lowest-achieving students.

Snapshot of performance trends in MONTENEGRO

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	392*	399*	412
PISA 2009	408*	403*	401*
PISA 2012	422	410*	410
PISA 2015	427	418*	411
PISA 2018	421	430	415
Average 3-year trend in mean performance	+7.7*	+7.6*	+1.7
Short-term change in mean performance (2015 to 2018)	-5.8	+11.7*	+3.9
Overall performance trajectory	positive, but flattening (less positive over more recent years)	increasingly positive	U-shaped (more positive over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.2	+0.7*	+0.1
Percentage-point change in low-achieving students (below Level 2)	-5.1	-10.5*	-2.0
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+8.0*	+7.8*	+2.6
Average trend amongst the lowest-achieving students (10th percentile)	+8.2*	+7.8*	+1.0
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Reading and mathematics performance in Montenegro improved since its first participation in PISA in 2006. In reading, most of the improvement occurred in earlier cycles, whereas in mathematics, most of the improvement was observed over the most recent period. In 2018, science performance returned to 2006 levels after an initial slump. Similar trends were observed across the performance distribution: improvements amongst the highest-performing students and amongst the lowest-performing students were close to those observed on average. In mathematics, these improvements resulted in a reduction in the share of low achievers (students scoring below Level 2) of more than 10 percentage points since 2012.

Snapshot of performance trends in the NETHERLANDS

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	513*	538*	
PISA 2006	507*	531*	525*
PISA 2009	508*	526	522*
PISA 2012	511*	523	522*
PISA 2015	503*	512	509
PISA 2018	485	519	503
Average 3-year trend in mean performance	-4.3*	-4.2*	-5.6*
Short-term change in mean performance (2015 to 2018)	-18.2*	+7.0	-5.2
Overall performance trajectory	increasingly negative	steadily negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.7	-0.8	-2.5
Percentage-point change in low-achieving students (below Level 2)	+9.8*	+1.0	+7.1*
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.6	-4.1*	-2.9*
Average trend amongst the lowest-achieving students (10th percentile)	-9.0*	-5.2*	-8.5*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In the Netherlands, mean performance in reading in 2018 was below the level observed in any previous assessment, while mean performance in mathematics and science remained closer to the level observed in 2015. However, when considering all comparable assessments, the long-term trajectory was clearly negative in mathematics and science too. In reading, no decline was observed amongst the highest-performing students, but rapid declines were observed amongst the lowest-achieving students; in science, performance declined amongst the highest-achieving students too, but more so amongst the lowest-achieving students. In mathematics, trends were similar across high- and low-achieving students. The proportion of students scoring at Level 5 or 6 in PISA remained stable in all three subjects, compared to the last assessment in which each subject was the major focus. However, the proportion of students scoring below Level 2 grew by almost 10 percentage points in reading (compared to 2009) and by 7 percentage points in science (compared to 2006).

Snapshot of performance trends in NEW ZEALAND

Mean performance	Reading	Mathematics	Science
PISA 2000	529*		
PISA 2003	522	523*	
PISA 2006	521*	522*	530*
PISA 2009	521*	519*	532*
PISA 2012	512	500	516
PISA 2015	509	495	513
PISA 2018	506	494	508
Average 3-year trend in mean performance	-3.7*	-7.0*	-6.2*
Short-term change in mean performance (2015 to 2018)	-3.5	-0.7	-4.8
Overall performance trajectory	steadily negative	steadily negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-2.6	-3.4*	-6.3*
Percentage-point change in low-achieving students (below Level 2)	+4.6*	-0.9	+4.3*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-3.3*	-7.9*	-7.1*
Average trend amongst the lowest-achieving students (10th percentile)	-3.2*	-6.0*	-5.1*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in New Zealand has been steadily declining in reading (2000-2018), mathematics (2003-2018) and science (2006-2018) from initially high levels of performance; it has been declining in science too, at least since 2012. In reading, more rapid declines were observed amongst the country's lowest-achieving students; in mathematics and science, performance declined to a similar extent at the top and the bottom of the performance distribution, as well as on average.

The proportion of top-performing students (scoring at Level 5 or 6) remained stable in reading (between 2009 and 2018), but decreased in mathematics (between 2012 and 2018) and in science (between 2006 and 2018). Meanwhile, the proportion of low-achieving students (scoring below Level 2) increased in reading and science.

Snapshot of performance trends in the REPUBLIC OF NORTH MACEDONIA

Mean performance	Reading	Mathematics	Science
PISA 2000	373*		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	m	m	m
PISA 2012	m	m	m
PISA 2015	352*	371*	384*
PISA 2018	393	394	413
Average 3-year trend in mean performance	+1.1	+23.3*	+28.7*
Short-term change in mean performance (2015 to 2018)	+40.9*	+23.1*	+29.4*
Overall performance trajectory	stable	improving	improving
Proficiency levels	**Reading (2015 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.7	-0.8	-2.5
Percentage-point change in low-achieving students (below Level 2)	+9.8*	+1.0	+7.1*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2015 to 2018)**	**Science (2015 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.6	-4.1*	-2.9*
Average trend amongst the lowest-achieving students (10th percentile)	-9.0*	-5.2*	-8.5*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Students in the Republic of North Macedonia (hereafter, North Macedonia) improved significantly in all three subjects since 2015. While performance is still significantly below the OECD average in reading, mathematics and science, the percentage of low performers in each subject shrank by at least 9 percentage points. Improvements were observed throughout the performance distribution, as the lowest- and highest-achieving students improved their proficiency between 2015 and 2018. The highest- and lowest-performing students in mathematics saw similar improvements in performance, while the highest-performing students in science improved significantly more than the lowest-performing students.

North Macedonia also participated in the reading assessment in PISA 2000; if these results were taken into account, mean reading performance in North Macedonia would be classified as stable.

Snapshot of performance trends in NORWAY

Mean performance	Reading	Mathematics	Science
PISA 2000	505		
PISA 2003	500	495	
PISA 2006	484*	490*	487
PISA 2009	503	498	500
PISA 2012	504	489*	495
PISA 2015	513*	502	498*
PISA 2018	499	501	490
Average 3-year trend in mean performance	+1.0	+1.5	+0.6
Short-term change in mean performance (2015 to 2018)	-13.7*	-0.8	-8.1*
Overall performance trajectory	flat	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.9*	+2.8*	+0.7
Percentage-point change in low-achieving students (below Level 2)	+4.3*	-3.4*	-0.2
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+1.3	+0.5	+1.9
Average trend amongst the lowest-achieving students (10th percentile)	+1.4	+1.8	-2.7
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In 2018, Norway's performance in PISA lay below PISA 2015 performance in reading and science. However, when trends were assessed over a longer period, no clear direction of change (neither positive, nor negative) could be determined in any subject. PISA 2018 results were close to the average performance across PISA assessments for the country. Trends over this longer period were similar at the top and at the bottom of the performance distribution.

At least over the more recent period (2009-2018), performance trends in Norway were influenced by the concurrent increase in the proportion of immigrant students who tended to score below non-immigrant students. It could be estimated that, if the student population in 2009 had had the same demographic profile as the population in 2018, the average score in reading would have been 497 points (Table I.B1.40). In reality, the average score observed in 2009 was 503 points (Table I.B1.10). The (non-significant) decline in mean performance between 2009 and 2018 could therefore be entirely explained by the changing demographic composition of the student population.

Snapshot of performance trends in PANAMA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	371	360	376
PISA 2012	m	m	m
PISA 2015	m	m	m
PISA 2018	377	353	365
Average 3-year trend in mean performance	+2.1	-2.3	-3.8
Short-term change in mean performance (2015 to 2018)	m	m	m
Overall performance trajectory	stable	stable	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.3	m	-0.1
Percentage-point change in low-achieving students (below Level 2)	-0.9	m	+6.2
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-2.8	-4.0	-5.5
Average trend amongst the lowest-achieving students (10th percentile)	+6.3	-2.1	-0.4
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Panama participated in PISA for the second time in 2018, after first participating in 2009. PISA 2018 results reflected broadly similar performance in all three subjects (reading, mathematics and science) as was observed in 2009.

Snapshot of performance trends in PERU

Mean performance	Reading	Mathematics	Science
PISA 2000	327*		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	370*	365*	369*
PISA 2012	384*	368*	373*
PISA 2015	398	387*	397
PISA 2018	401	400	404
Average 3-year trend in mean performance	+13.5*	+12.2*	+12.8*
Short-term change in mean performance (2015 to 2018)	+3.0	+13.3*	+7.5
Overall performance trajectory	positive, but flattening (less positive over more recent years)	improving	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.3	+0.3	+0.0
Percentage-point change in low-achieving students (below Level 2)	-10.5*	-14.2*	-13.8*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+12.5*	+10.8*	+10.3*
Average trend amongst the lowest-achieving students (10th percentile)	+14.6*	+14.5*	+17.3*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Over the 2009-2018 period, mean performance in Peru improved from initially low levels in all three subjects (reading, mathematics and science). The improvement in reading performance is even more pronounced when considering PISA 2000 results. Improvements were observed amongst the country's highest-achieving and lowest-achieving students. In mathematics, a significant improvement was also observed over the most recent period (2015-2018).

Snapshot of performance trends in POLAND

Mean performance	Reading	Mathematics	Science
PISA 2000	479*		
PISA 2003	497	490*	
PISA 2006	508	495*	498*
PISA 2009	500*	495*	508
PISA 2012	518	518	526*
PISA 2015	506	504*	501*
PISA 2018	512	516	511
Average 3-year trend in mean performance	+4.5*	+5.1*	+2.1
Short-term change in mean performance (2015 to 2018)	+6.2	+11.2*	+9.6*
Overall performance trajectory	positive, but flattening (less positive over more recent years)	steadily positive	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+5.0*	-1.0	+2.5*
Percentage-point change in low-achieving students (below Level 2)	-0.4	+0.3	-3.1*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+3.3*	+4.7*	+3.0
Average trend amongst the lowest-achieving students (10th percentile)	+6.4*	+4.4*	+1.0
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Poland, PISA 2018 performance was about 10 points higher than in 2015 in mathematics and science, and close to the level observed in 2012 in reading and mathematics. Over the longer period, the direction of change in mean performance was clearly positive in reading (2000-2018) and mathematics (2003-2018). In science, no clear direction of change could be determined, because PISA 2018 results were higher than results observed in 2006 and 2015 (when science was the focus of the assessment), but remained below those observed in 2012.

Between 2009 and 2018, the proportion of top-performing students in reading (students scoring at Level 5 or 6) increased by 5 percentage points.

Snapshot of performance trends in PORTUGAL

Mean performance	Reading	Mathematics	Science
PISA 2000	470*		
PISA 2003	478	466*	
PISA 2006	472*	466*	474*
PISA 2009	489	487	493
PISA 2012	488	487	489
PISA 2015	498	492	501*
PISA 2018	492	492	492
Average 3-year trend in mean performance	+4.3*	+6.0*	+4.3*
Short-term change in mean performance (2015 to 2018)	-6.3	+0.9	-9.4*
Overall performance trajectory	steadily positive	positive, but flattening (less positive over more recent years)	positive, but flattening (less positive over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.5*	+1.0	+2.5*
Percentage-point change in low-achieving students (below Level 2)	+2.6	-1.6	-4.9*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.1*	+7.8*	+6.0*
Average trend amongst the lowest-achieving students (10th percentile)	+5.2*	+2.6	+1.7
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Portugal, mean performance in reading, mathematics and science improved since 2000, 2003 and 2006. In reading and mathematics, mean performance in 2018 was close to the level observed over the period 2009-2015; in science, mean performance in 2018 was below that of 2015, and returned close to the level observed in 2009 and 2012.

Snapshot of performance trends in QATAR

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	312*	318*	349*
PISA 2009	372*	368*	379*
PISA 2012	388*	376*	384*
PISA 2015	402	402*	418
PISA 2018	407	414	419
Average 3-year trend in mean performance	+21.9*	+22.6*	+17.9*
Short-term change in mean performance (2015 to 2018)	+5.2	+11.8*	+1.5
Overall performance trajectory	positive, but flattening (less positive over more recent years)	positive, but flattening (less positive over more recent years)	positive, but flattening (less positive over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.9*	+0.9*	+1.9*
Percentage-point change in low-achieving students (below Level 2)	-12.6*	-15.9*	-30.7*
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+20.9*	+23.9*	+22.2*
Average trend amongst the lowest-achieving students (10th percentile)	+19.3*	+18.1*	+11.3*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Performance in reading, mathematics and science in Qatar improved at one of the most rapid rates, and from initially low levels, throughout the country's participation in PISA. As a result, in all subjects, the share of low-achieving students (those who scored below Level 2) shrank and the share of top-performing students (those who scored at Level 5 or 6) increased.

Over the most recent period (2009-2018), about one-third of the improvement in reading performance (i.e. 13 of 35 score points) could be attributed to changes in the composition of the student population in Qatar, with significant increases in the share of foreign-born students, who tended to score higher than non-immigrant students (Table I.B1.40).

Snapshot of performance trends in ROMANIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	396*	415*	418
PISA 2009	424	427	428
PISA 2012	438	445*	439
PISA 2015	434	444*	435
PISA 2018	428	430	426
Average 3-year trend in mean performance	+7.2*	+4.7*	+2.1
Short-term change in mean performance (2015 to 2018)	-5.9	-14.0*	-9.1
Overall performance trajectory	positive, but flattening (less positive over more recent years)	positive, but flattening (less positive over more recent years)	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.7*	-0.0	+0.5
Percentage-point change in low-achieving students (below Level 2)	+0.4	+5.7	-3.0
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+10.1*	+8.8*	+4.5*
Average trend amongst the lowest-achieving students (10th percentile)	+5.4*	+1.2	+0.1
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	widening gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean reading performance in Romania was higher than in 2006, when the country first participated in PISA, but not statistically significantly different from any result since then. Mean mathematics performance in 2018 was significantly lower than in 2015, reversing some of the gains observed between 2006 and 2015 whereas science performance returned, in 2018, close to the level observed in 2006 or 2009.

In mathematics and science, students at different levels in the performance distribution followed distinct trends, and gaps in performance widened.

Snapshot of performance trends in the RUSSIAN FEDERATION

Mean performance	Reading	Mathematics	Science
PISA 2000	462*		
PISA 2003	442*	468*	
PISA 2006	440*	476*	479
PISA 2009	459*	468*	478
PISA 2012	475	482	486
PISA 2015	495*	494	487*
PISA 2018	479	488	478
Average 3-year trend in mean performance	+6.8*	+4.7*	+0.5
Short-term change in mean performance (2015 to 2018)	-16.1*	-6.3	-8.9*
Overall performance trajectory	increasingly positive	steadily positive	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.3*	+0.3	-1.1
Percentage-point change in low-achieving students (below Level 2)	-5.3*	-2.3	-1.0
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+6.7*	+2.8	-1.9
Average trend amongst the lowest-achieving students (10th percentile)	+7.7*	+5.8*	+2.5
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In the Russian Federation, performance in PISA 2018 was close to the level observed in 2012 in all three subjects, although it lay below PISA 2015 performance in reading and mathematics. Over a longer period of time, and when taking into account results from all years, the overall direction of mean performance trends was positive in reading (over the 2000-2018 period) and in mathematics (over the 2003-2018 period), while no overall direction of change can be determined in science.

In science, a more positive trend was observed amongst the country's lowest-achieving students than amongst the country's highest-achieving students. In mathematics, performance improved both at the top and at the bottom of the distribution, but more so amongst the lowest-achieving students (at the bottom). As a result, performance gaps in these two subjects narrowed.

Snapshot of performance trends in SERBIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	401*	435*	436
PISA 2009	442	442	443
PISA 2012	446	449	445
PISA 2015	m	m	m
PISA 2018	439	448	440
Average 3-year trend in mean performance	+7.7*	+3.0*	+0.7
Short-term change in mean performance (2015 to 2018)	m	m	m
Overall performance trajectory	improving	improving	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.7*	+0.7	+0.8*
Percentage-point change in low-achieving students (below Level 2)	+4.9	+0.8	-0.2
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+11.3*	+5.6*	+4.4*
Average trend amongst the lowest-achieving students (10th percentile)	+4.8*	+1.2	-1.8
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Serbia, mean performance in reading and mathematics improved since the country first participated in PISA in 2006; performance in science remained stable, on average. Across all three subjects, improvements were more marked amongst the highest-achieving students, and a widening of performance gaps was observed. The percentage of students scoring at the highest levels of proficiency increased, particularly in reading (+1.7 percentage points since 2009) and in science (+0.8 of a percentage point since 2006).

Snapshot of performance trends in SINGAPORE

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	526*	562	542*
PISA 2012	542	573	551
PISA 2015	535*	564	556
PISA 2018	549	569	551
Average 3-year trend in mean performance	+6.4*	+1.1	+3.2*
Short-term change in mean performance (2015 to 2018)	+14.4*	+4.8	-4.6
Overall performance trajectory	improving	stable	improving
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+10.1*	-3.1	+0.8
Percentage-point change in low-achieving students (below Level 2)	-1.2	-1.1	-2.5*
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+9.5*	-5.1*	-0.9
Average trend amongst the lowest-achieving students (10th percentile)	+0.3	+5.9*	+4.4*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in Singapore improved significantly in reading, both over the longer period (2009-2018) and between 2015 and 2018. Mean mathematics performance in 2018 stood close to the average level observed over previous assessments (2009-2015), while mean performance in science improved between 2009 and 2018.

Performance in reading improved, particularly amongst the country's highest-achieving students. Between 2009 and 2018 the proportion of 15-year-old students scoring at Level 5 or 6 on the PISA scale increased by 10 percentage points; meanwhile, the proportion of low-achieving students in reading remained stable. In science, by contrast, improvements in mean performance were driven by improvements amongst the lowest-achieving students: the proportion of low-achievers in science (those scoring below Level 2) shrank by 2.5 percentage points.

Snapshot of performance trends in the SLOVAK REPUBLIC

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	469	498*	
PISA 2006	466	492	488*
PISA 2009	477*	497*	490*
PISA 2012	463	482	471
PISA 2015	453	475*	461
PISA 2018	458	486	464
Average 3-year trend in mean performance	-3.2*	-3.6*	-7.8*
Short-term change in mean performance (2015 to 2018)	+5.5	+10.9*	+3.3
Overall performance trajectory	steadily negative	steadily negative	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+0.2	-0.2	-2.1*
Percentage-point change in low-achieving students (below Level 2)	+9.2*	-2.4	+9.1*
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-0.8	-2.8*	-6.2*
Average trend amongst the lowest-achieving students (10th percentile)	-5.4*	-6.1*	-10.0*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in science and, to a lesser extent, in reading and mathematics, has declined steadily since the Slovak Republic's first participation in PISA. In science, mean performance in 2018 was roughly 25 score points below what it was in 2006 and 2009. In mathematics, performance in 2015 was particularly poor, but PISA 2018 results marked a return to a level similar to that observed in 2012. In reading, the decline was the mildest. Amongst the lowest-achieving students, performance declined, on average, by 5.4 score points every 3 years, whereas amongst the highest-performing students, performance remained stable. In reading, the proportion of low-achieving students (students scoring below Level 2) grew by about 9 percentage points between 2009 and 2018; a similarly large increase was observed in science between 2009 and 2018.

Snapshot of performance trends in SLOVENIA

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	494	504	519*
PISA 2009	483*	501	512
PISA 2012	481*	501*	514
PISA 2015	505*	510	513*
PISA 2018	495	509	507
Average 3-year trend in mean performance	+2.4	+1.8	-2.2*
Short-term change in mean performance (2015 to 2018)	-9.9*	-1.0	-5.9*
Overall performance trajectory	U-shaped (more positive over more recent years)	U-shaped (more positive over more recent years)	steadily negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.2*	-0.1	-5.6*
Percentage-point change in low-achieving students (below Level 2)	-3.3*	-3.7*	+0.7
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+4.3*	-0.8	-5.0*
Average trend amongst the lowest-achieving students (10th percentile)	+1.3	+2.0	-0.2
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Slovenia, PISA 2018 results in reading and mathematics lay close to their historic average. Some significant differences were observed when comparing PISA 2018 results to those of earlier years. In particular, PISA 2018 reading performance was lower, on average, than in 2015, but higher than in 2009 or 2012; and PISA 2018 mathematics performance was higher than in 2012. However, over the full 2006-2018 period, and when taking into account results from all years, no significant improving or declining trend could be determined. In science, a mild negative trend was observed, and performance was lower, on average, than in 2015 and in 2006.

Between 2009 and 2018, improvements in reading performance appeared to be particularly strong amongst the highest-achieving students; and the proportion of students scoring at Level 5 or 6 in PISA (top-performing students) increased by 3.2 percentage points. In science, by contrast, between 2006 and 2018, performance amongst the highest-achieving students appeared to decline faster than amongst the lowest-achieving students. The proportion of top-performing students in science shrank by 5.6 percentage points over this period.

Snapshot of performance trends in SPAIN

Mean performance	Reading	Mathematics	Science
PISA 2000	493*		
PISA 2003	481	485	
PISA 2006	461*	480	488
PISA 2009	481	483	488
PISA 2012	488*	484	496*
PISA 2015	496*	486	493*
PISA 2018	m	481	483
Average 3-year trend in mean performance	+1.6	+0.0	-0.5
Short-term change in mean performance (2015 to 2018)	m	-4.5	-9.5*
Overall performance trajectory	U-shaped (more positive over more recent years)	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	m	-0.7	-0.7
Percentage-point change in low-achieving students (below Level 2)	m	+1.1	+1.6
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	m	-0.6	-0.2
Average trend amongst the lowest-achieving students (10th percentile)	m	+0.4	-0.9
Gap in learning outcomes between the highest- and lowest-achieving students	m	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Spain's data met PISA 2018 Technical Standards. However, some data show implausible student-response behaviour. Consequently, at the time of publication of this report, comparability of Spain's results in reading cannot be assured (see Annex A9). PISA 2018 reading results for Spain are therefore not published in this report

Mean mathematics performance remained stable, around a flat trend line, throughout the country's participation in PISA (including PISA 2018). Mean performance in science declined between 2015 and 2018 by 9.5 score points. Despite the recent decline in science performance, over a longer period, and when taking into account results from all years, no significant improving or declining trend could be determined, in any subject.

Snapshot of performance trends in SWEDEN

Mean performance	Reading	Mathematics	Science
PISA 2000	516		
PISA 2003	514	509	
PISA 2006	507	502	503
PISA 2009	497	494	495
PISA 2012	483*	478*	485*
PISA 2015	500	494	493
PISA 2018	506	502	499
Average 3-year trend in mean performance	-3.0*	-2.1	-1.0
Short-term change in mean performance (2015 to 2018)	+5.6	+8.5	+6.0
Overall performance trajectory	negative, but flattening (less negative over more recent years)	U-shaped (more positive over more recent years)	U-shaped (more positive over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+4.2*	+4.6*	+0.4
Percentage-point change in low-achieving students (below Level 2)	+1.0	-8.3*	+2.6
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.1	-2.9*	+0.5
Average trend amongst the lowest-achieving students (10th percentile)	-6.4*	-2.0	-3.8*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

After a rapid decline until 2012, mean reading, mathematics and science performance in Sweden recovered fully or almost fully between 2012 and 2018, returning to a level similar to that observed in the early PISA assessments. In mathematics, for example, mean performance in 2018 lay more than 20 points above the PISA 2012 mean score. Between 2012 and 2018, the proportion of low-achieving students (scoring below Level 2) shrank by 8 percentage points and, at the same time, the proportion of top-performing students (scoring at Level 5 or 6) grew by about 5 percentage points. In reading and science, however, performance gaps widened over the long term. While no overall change could be determined amongst the highest-achieving students, performance amongst the lowest-achieving students tended to decline, particularly in reading.

Sweden's improvement in mean performance since PISA 2012 was observed over a period of rapid increase in the proportion of immigrant students, who tended to score below non-immigrant students. It could be estimated that, if the student population in 2009 had had the same demographic profile as the population in 2018, the average score in reading would have been nine points lower than what was observed that year (Tables I.B1.10 and I.B1.40) – and the recent trends would have been even more positive. The widening gap in reading performance between the highest- and lowest-achieving students also seemed to be at least partly related to growing shares of immigrant students (Tables I.B1.10 and I.B1.40).

The massive inflow of immigrants in the most recent period, however, also led to an increase in student exclusion rates. In 2018, about 11% of 15-year-old students were excluded from the PISA test – the highest rate amongst all participating countries/economies (Table I.A2.1). While limited information is available about excluded students, this increase is most likely the consequence of the large (and temporary) increase, between 2015 and 2018, of recently arrived immigrants in the school system.

Snapshot of performance trends in SWITZERLAND

Mean performance	Reading	Mathematics	Science
PISA 2000	494		
PISA 2003	499	527*	
PISA 2006	499*	530*	512*
PISA 2009	501*	534*	517*
PISA 2012	509*	531*	515*
PISA 2015	492	521	506*
PISA 2018	484	515	495
Average 3-year trend in mean performance	-1.3	-2.5*	-4.4*
Short-term change in mean performance (2015 to 2018)	-8.3	-5.9	-10.2*
Overall performance trajectory	hump-shaped (more negative over more recent years)	increasingly negative	increasingly negative
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.0	-4.4*	-2.7*
Percentage-point change in low-achieving students (below Level 2)	+6.8*	+4.4*	+4.2*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-0.5	-3.4*	-3.3
Average trend amongst the lowest-achieving students (10th percentile)	-1.7	-1.1	-3.9*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In 2018, mean performance in reading, mathematics and science in Switzerland was significantly below mean performance in PISA 2006, 2009 or 2012. The decline in performance was particularly marked since 2012. Overall trends followed similar trajectories at the top and bottom of the performance distribution. In mathematics, for example, the proportion of top-performing students (scoring at Level 5 or 6) shrank by 4.4 percentage points between 2012 and 2018, and the proportion of low-achieving students (scoring below Level 2) increased by a similar amount.

Between 2009 and 2018, the proportion of 15-year-old students with an immigrant background in Switzerland increased by 10 percentage points, one of the largest increases amongst OECD countries (Table II.B1.9.9 in *PISA 2018 Results (Volume II): Where All Students Can Succeed*). Because in Switzerland, in 2009 as well as in 2018, immigrant students scored about 50 points below non-immigrant students in reading (Table II.B1.9.10), the change in the proportion of immigrant and non-immigrant students alone could account for about five points (i.e. 10% of 50 points), or roughly one-third of the 17-point difference in mean reading scores between PISA 2009 and PISA 2018 (see also Tables I.B1.40-I.B1.42 for mean performance trends that account for demographic changes in the student population).

Snapshot of performance trends in CHINESE TAIPEI

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	496	549*	532*
PISA 2009	495	543*	520
PISA 2012	523*	560*	523
PISA 2015	497	542*	532*
PISA 2018	503	531	516
Average 3-year trend in mean performance	+1.5	-3.8*	-2.2
Short-term change in mean performance (2015 to 2018)	+5.5	-11.2*	-16.6*
Overall performance trajectory	hump-shaped (more negative over more recent years)	increasingly negative	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+5.7*	-14.0*	-3.0*
Percentage-point change in low-achieving students (below Level 2)	+2.2	+1.1	+3.5*
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+7.4*	-5.2*	+0.5
Average trend amongst the lowest-achieving students (10th percentile)	-3.7	-2.4	-4.6*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	stable gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Performance in Chinese Taipei has fluctuated since its first participation in PISA in 2006. The trajectory of mean performance in science could be classified as flat overall, although that masks relative highs in performance in 2006 and 2012 and relative lows in 2009, 2015 and 2018.

The trajectory of mean performance in reading can be described as hump-shaped, primarily due to Chinese Taipei's high performance in 2012. Performance in all other years was statistically similar to that observed in 2018. The gap in performance between the highest- and lowest-achieving students widened, primarily due to the highest-achieving students performing better over time (by 7.4 score points every 3 years). Between 2008 and 2018, the proportion of students who scored at Level 5 or 6 in reading grew by about 6 percentage points, but the proportion of low achievers (scoring below Level 2) did not decrease.

The trajectory was more negative in mathematics, where PISA 2018 results were significantly lower than in any previous year, and particularly compared to 2012 results (a decline of 29 score points), the last time mathematics was the focus of the assessment. The highest-achieving students performed worse in mathematics over time, declining 5.2 score points every 3 years on average over the 2006 to 2018 period; and the proportion of top-performing students (scoring at Level 5 or 6) shrank by 14 percentage points between 2012 and 2018. Nevertheless, mean performance in mathematics remained well above the OECD average.

Snapshot of performance trends in THAILAND

Mean performance	Reading	Mathematics	Science
PISA 2000	431*		
PISA 2003	420*	417	
PISA 2006	417*	417	421
PISA 2009	421*	419	425
PISA 2012	441*	427	444*
PISA 2015	409*	415	421
PISA 2018	393	419	426
Average 3-year trend in mean performance	-4.1*	+0.3	+0.6
Short-term change in mean performance (2015 to 2018)	-16.2*	+3.1	+4.5
Overall performance trajectory	increasingly negative	flat	hump-shaped (more negative over more recent years)
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.1	-0.3	+0.3
Percentage-point change in low-achieving students (below Level 2)	+16.7*	+3.0	-1.6
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-2.6	+1.4	+2.3
Average trend amongst the lowest-achieving students (10th percentile)	-4.1*	-1.1	-0.5
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	widening gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Thailand's mean reading performance in PISA 2018 was lower than in any previous assessment, and 16 points lower than in PISA 2015. In mathematics, by contrast, mean performance remained stable, around a flat trend line, over the entire period (2003-2018). Performance in science also appeared stable; only PISA 2012 results differ significantly from PISA 2018 results.

The negative trend in reading resulted in an increase of 16.7 percentage points, between 2009 and 2018, in the proportion of students scoring below Level 2 (low achievers).

Snapshot of performance trends in TURKEY

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	441*	423*	
PISA 2006	447*	424*	424*
PISA 2009	464	445	454*
PISA 2012	475	448	463
PISA 2015	428*	420*	425*
PISA 2018	466	454	468
Average 3-year trend in mean performance	+2.2	+4.1*	+6.1*
Short-term change in mean performance (2015 to 2018)	+37.3*	+33.1*	+42.8*
Overall performance trajectory	hump-shaped (more negative over more recent years)	steadily positive	steadily positive
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+1.5*	-1.1	+1.5*
Percentage-point change in low-achieving students (below Level 2)	+1.6	-5.3	-21.4*
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.7	-0.2	+5.1*
Average trend amongst the lowest-achieving students (10th percentile)	+3.4*	+6.3*	+4.8*
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	narrowing gap	stable gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Turkey's mean performance in PISA 2018, in all three subjects, was not significantly different from that observed in 2009 or 2012 and was significantly higher than the level observed in 2003 and 2006. When considering results from all years, it is clear that PISA 2015 results – which were considerably lower – were anomalous, and neither the decline between 2012 and 2015, nor the recovery between 2015 and 2018, reflect the long-term trajectory. Overall, this trajectory is clearly positive in mathematics (over the 2003-2018 period) and in science (2006-2018). In mathematics, improvements were more pronounced at the bottom of the performance distribution, amongst the lowest-achieving students, who caught up to the higher-performing students.

These performance trends were observed over a period of rapid expansion of secondary education. Between 2003 and 2018, Turkey added more than 400 000 students to the total population of 15-year-olds eligible to participate in PISA; the proportion of 15-year-olds who were covered by PISA samples more than doubled, from about 36% in 2003 to 73% in 2018 (Table I.A2.1). It is likely that this expansion in education opportunities dampened a more positive underlying trend in student performance. Indeed, a simulation that assumes that the top-scoring 25% of 15-year-olds were eligible to take the test in any given year shows a positive trend amongst this population in mathematics (since 2003) and science (since 2006) (Figure I.9.5).

Snapshot of performance trends in the UNITED ARAB EMIRATES

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	m	m	m
PISA 2009	431	421*	438
PISA 2012	442*	434	448*
PISA 2015	434	427	437
PISA 2018	432	435	434
Average 3-year trend in mean performance	-0.7	+3.7	-2.5
Short-term change in mean performance (2015 to 2018)	-1.8	+7.5	-3.1
Overall performance trajectory	stable	stable	stable
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2009 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+2.5*	+2.0*	+0.7*
Percentage-point change in low-achieving students (below Level 2)	+3.1	-0.8	+3.6
Variation in performance	**Reading (2009 to 2018)**	**Mathematics (2009 to 2018)**	**Science (2009 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+8.9*	+10.0*	+3.0
Average trend amongst the lowest-achieving students (10th percentile)	-8.1*	-3.7	-6.5*
Gap in learning outcomes between the highest- and lowest-achieving students	widening gap	widening gap	widening gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Since the United Arab Emirates first participated in PISA in 2009, mean performance across all three subjects remained largely stable. Mean performance in mathematics fluctuated, but only over a range of less than 15 score points. This apparent stability masks changes in the performance distribution, however. In all three subjects, the highest-achieving students either improved their performance (by up to 10 score points every 3 years in mathematics) or saw no significant change in their performance. The lowest-achieving students either saw a decline in their performance (by up to 8.1 score points every 3 years in reading) or saw no significant change. Since 2009, the gap between the highest- and lowest-achieving students widened in all three subjects.

Snapshot of performance trends in the UNITED KINGDOM

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	m	m	
PISA 2006	495	495	515*
PISA 2009	494*	492	514
PISA 2012	499	494	514
PISA 2015	498	492*	509
PISA 2018	504	502	505
Average 3-year trend in mean performance	+2.1	+1.3	-2.4
Short-term change in mean performance (2015 to 2018)	+6.0	+9.3*	-4.6
Overall performance trajectory	flat	U-shaped (more positive over more recent years)	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.4*	+1.0	-4.1*
Percentage-point change in low-achieving students (below Level 2)	-1.1	-2.6	+0.7
Variation in performance	**Reading (2006 to 2018)**	**Mathematics (2006 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+2.9	+1.9	-4.2*
Average trend amongst the lowest-achieving students (10th percentile)	+2.9	-0.8	-1.0
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in reading and science in the United Kingdom remained stable since 2006, with no significant improvement or decline. This apparently stability hides changes in the performance of high- and low-scoring students. There was a 3.4 percentage-point increase in the share of top performers in reading between 2009 and 2018 but a 4.1 percentage-point decrease in the percentage of top performers in science between 2006 and 2018. Mean performance in mathematics was mostly flat but with a significant 9 score-point improvement between 2015 and 2018. In 2018, for the first time, the United Kingdom performed statistically significantly above the OECD average in mathematics.

Snapshot of performance trends in the UNITED STATES

Mean performance	Reading	Mathematics	Science
PISA 2000	504		
PISA 2003	495	483	
PISA 2006	m	474	489*
PISA 2009	500	487	502
PISA 2012	498	481	497
PISA 2015	497	470	496
PISA 2018	505	478	502
Average 3-year trend in mean performance	+0.2	-1.2	+2.1
Short-term change in mean performance (2015 to 2018)	+8.4	+8.6	+6.1
Overall performance trajectory	flat	flat	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	+3.7*	-0.5	+0.0
Percentage-point change in low-achieving students (below Level 2)	+1.6	+1.3	-5.7*
Variation in performance	**Reading (2000 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	+0.4	-2.3	-0.2
Average trend amongst the lowest-achieving students (10th percentile)	+0.2	-0.1	+3.6
Gap in learning outcomes between the highest- and lowest-achieving students	stable gap	stable gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

Mean performance in reading, mathematics and science in the United States remained about the same in every PISA assessment, with no significant improvement or decline. Only science performance in 2006 was significantly below the 2018 mean score, but even in science, performance has followed a flat trajectory since 2009.

Nevertheless, in reading, the share of 15-year-old students scoring at Level 5 or 6 (top performers) increased by almost 4 percentage points between 2009 and 2018, to 13.5%. In science, some improvements were observed amongst the lowest-achieving students, and the gap between the lowest- and the highest-achieving students narrowed. The share of 15-year-old students scoring below Level 2 proficiency in science shrank by 5.7 percentage points between 2006 and 2018.

Snapshot of performance trends in URUGUAY

Mean performance	Reading	Mathematics	Science
PISA 2000	m		
PISA 2003	434	422	
PISA 2006	413*	427	428
PISA 2009	426	427	427
PISA 2012	411*	409	416
PISA 2015	437	418	435*
PISA 2018	427	418	426
Average 3-year trend in mean performance	+0.6	-2.0	+0.4
Short-term change in mean performance (2015 to 2018)	-9.5	-0.3	-9.6*
Overall performance trajectory	U-shaped (more positive over more recent years)	flat	flat
Proficiency levels	**Reading (2009 to 2018)**	**Mathematics (2012 to 2018)**	**Science (2006 to 2018)**
Percentage-point change in top-performing students (Level 5 or 6)	-0.2	-0.3	-0.7*
Percentage-point change in low-achieving students (below Level 2)	-0.0	-5.1	+1.7
Variation in performance	**Reading (2003 to 2018)**	**Mathematics (2003 to 2018)**	**Science (2006 to 2018)**
Average trend amongst the highest-achieving students (90th percentile)	-5.8*	-5.4*	-1.9
Average trend amongst the lowest-achieving students (10th percentile)	+8.4*	+3.1*	+4.0*
Gap in learning outcomes between the highest- and lowest-achieving students	narrowing gap	narrowing gap	narrowing gap

* indicates statistically significant trends and changes, or mean-performance estimates that are significantly above or below PISA 2018 estimates.

Note: Differences between PISA 2003-2012 scores and PISA 2015-2018 scores in Uruguay may also reflect a different treatment of non-reached items (missing answers to items placed at the end of the test). See ANEP, INEEd and UDELAR (2019), *Informe del grupo técnico para la comparabilidad de los resultados de PISA 2015 con anteriores ciclos de la evaluación en Uruguay*.

Source: OECD, PISA 2018 Database, Tables I.B1.7–I.B1.15 and I.B1.28–I.B1.30.

In Uruguay, PISA 2018 performance in all three subjects was close to the levels observed in its first participation in 2003 (or 2006 for science). The poorest performance in all three subjects was observed in 2012, after which performance returned to previous levels. A peak in reading and science performance was observed in 2015. However, this description hides changes in the performance distribution over time. In all three subjects, the performance of the lowest-achieving students improved since Uruguay first participated in PISA, while there was either a drop or a lack of significant change in performance amongst the highest-achieving students. These trends have resulted in a narrowing of the gap between the highest- and lowest-achieving students over the period.

Uruguay increased coverage of its 15-year-old population since 2003: in 2018, PISA covered 78% of the country's 15-year-olds, compared to 63% in 2003 and 2009. Greater enrolment often involves the inclusion of relatively weaker students; thus maintaining performance at the same level while enrolment increases is often a sign of improvement in the education system. On the assumption that 15-year-olds who were excluded would have performed below the median if they had sat the PISA assessment, Uruguay saw an improvement in the performance of the median 15-year-old by 15 score points in reading, 7 score points in mathematics and 7 score points in science every three years (Tables I.B1.34-I.B1.36).

ANNEX E

The development and implementation of PISA: A collaborative effort

PISA is a collaborative effort, bringing together experts from the participating countries, steered jointly by their governments on the basis of shared, policy-driven interests.

A PISA Governing Board, on which each country is represented, determines the policy priorities for PISA, in the context of OECD objectives, and oversees adherence to these priorities during the implementation of the programme. This includes setting priorities for the development of indicators, for establishing the assessment instruments, and for reporting the results.

Experts from participating countries also serve on working groups that are charged with linking policy objectives with the best internationally available technical expertise. By participating in these expert groups, countries ensure that the instruments are internationally valid and take into account the cultural and educational contexts in OECD member and partner countries and economies, that the assessment materials have strong measurement properties, and that the instruments emphasise authenticity and educational validity.

Through National Project Managers, participating countries and economies implement PISA at the national level subject to the agreed administration procedures. National Project Managers play a vital role in ensuring that the implementation of the survey is of high quality, and verify and evaluate the survey results, analyses, reports and publications.

The design and implementation of the surveys, within the framework established by the PISA Governing Board, is the responsibility of external contractors. For PISA 2018, the overall management of contractors and implementation was carried out by the Educational Testing Service (ETS) in the United States as the Core A contractor. Tasks under Core A also included instrument development, development of the computer platform, survey operations and meetings, scaling, analysis and data products. These tasks were implemented in co-operation with the following subcontractors; i) the University of Luxembourg for support with test development; ii) the Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe) at the University of Liège in Belgium for test development and coding training for open-response items; iii) the International Association for the Evaluation of Educational Achievement (IEA) in the Netherlands for the data management software; iv) Westat in the United States for survey operations; v) Deutsches Institut für Internationale Pädagogische Forschung (DIPF) in Germany, with co-operation from Statistics Canada, for the development of the questionnaires; and vi) HallStat SPRL in Belgium for the translation referee.

The remaining tasks related to the implementation of PISA 2018 were implemented through three additional contractors – Cores B to D. The development of the cognitive assessment frameworks for reading and global competence and of the framework for questionnaires was carried out by Pearson in the United Kingdom as the Core B contractor. Core C focused on sampling and was the responsibility of Westat in the United States in co-operation with the Australian Council for Educational Research (ACER) for the sampling software KeyQuest. Linguistic quality control and the development of the French source version for Core D were undertaken by cApStAn, who worked in collaboration with BranTra as a subcontractor.

The OECD Secretariat has overall managerial responsibility for the programme, monitors its implementation daily, acts as the secretariat for the PISA Governing Board, builds consensus among countries and serves as the interlocutor between the PISA Governing Board and the international Consortium charged with implementing the activities. The OECD Secretariat also produces the indicators and analyses and prepares the international reports and publications in co-operation with the PISA Consortium and in close consultation with member and partner countries and economies both at the policy level (PISA Governing Board) and at the level of implementation (National Project Managers).

PISA GOVERNING BOARD

(*Former PGB representative who was involved in PISA 2018)

Chair of the PISA Governing Board: Michele Bruniges

OECD Members and PISA Associates

Australia: Rick Persse, Rhyan Bloor* and Gabrielle Phillips*

Austria: Mark Német

Belgium: Isabelle Erauw and Geneviève Hindryckx

Brazil: Alexandre Ribeiro Pereira Lopes, Maria Helena Guimarães De Castro*, Maria Inês Fini* and José Francisco Soares*

Canada: Gilles Bérubé, Kathryn O'Grady, Pierre Brochu* and Tomasz Gluszynski*

Chile: Claudia Matus and Carolina Flores*

Czech Republic: Tomas Zatloukal

Denmark: Charlotte Rotbøll Sjøgreen, Hjalte Meilvang, Eyðun Gaard, Mette Hansen* and Frida Poulsen*

Estonia: Maie Kitsing

Finland: Tommi Karjalainen and Najat Ouakrim-Soivio*

France: Ronan Vourc'h, Thierry Rocher* and Bruno Trosseille*

Germany: Jens Fischer-Kottenstede, Katharina Koufen, Elfriede Ohrnberger and Martina Diedrich*

Greece: Ioannis Tsirmpas and Chryssa Sofianopoulou*

Hungary: Sándor Brassói

Iceland: Stefan Baldursson

Ireland: Rachel Perkins, Peter Archer* and Caroline McKeown*

Israel: Hagit Glickman

Italy: Roberto Ricci

Japan: Yu Kameoka and Akiko Ono*

Korea: Jimin Cho, Ji-Young Park, Dong-In Bae*, Inn-Soon Jung*, Sungsook Kim*, Myungae Lee*, Bu Ho Nam* and Jea Yun Park*

Latvia: Alona Babica and Liga Lejiņa*

Lithuania: Rita Dukynaite

Luxembourg: Amina Afif

Mexico: Andres Sanchez, Ana María Aceves Estrada*, Eduardo Backhoff Escudero* and Otto Granados Roldán*

Netherlands: Marjan Zandbergen

New Zealand: Craig Jones and Lisa Rodgers*

Norway: Marthe Akselsen and Anne-Berit Kavli*

Poland: Piotr Mikiewicz, Lidia Olak* and Jerzy Wiśniewski*

Portugal: Luís Pereira Dos Santos and Hélder Manuel Diniz De Sousa*

Slovak Republic: Romana Kanovska

Slovenia: Ksenija Bregar Golobic, Mojca Štraus and Andreja Barle Lakota*

Spain: Carmen Tovar Sánchez

Sweden: Ellen Almgren and Eva Lundgren*

Switzerland: Reto Furter, Camil Würgler, Vera Husfeldt* and Claudia Zahner Rossier*

Thailand: Sukit Limpijumnong, Nantawan Somsook and Supattra Pativisan*

Turkey: Sadri Şensoy and Kemal Bülbül*

United Kingdom: Lorna Bertrand, Keith Dryburgh and Jonathan Wright*

United States: Peggy Carr and Dana Kelly*

Observers (Partner economies)

Albania: Zamira Gjini

Argentina: María Angela Cortelezzi and Elena Duro*

Azerbaijan: Emin Amrullayev

Belarus: Aliaksandr Yakabchuk

Bosnia and Herzegovina: Maja Stojkic

Brunei Darussalam: Shamsiah Zuraini Kanchanawati Tajuddin, Hj Azman Bin Ahmad* and Hjh Romaizah Hj Mohd Salleh*

Bulgaria: Neda Oscar Kristanova

Beijing-Shanghai-Jiangsu-Zhejiang (China): Zhang Jin, Xiang Mingcan, Jun Fang*, Yanpin Hu* and Lin Shiliang*

Colombia: María Figueroa Cahnspeyer and Ximena Dueñas Herrera*

Costa Rica: Pablo José Mena Castillo, Melania Brenes Monge, Edgar Mora Altamirano* and Alicia Vargas Porras*

Croatia: Ines Elezovic and Michelle Bras Roth*

Dominican Republic: Ancell Scheker Mendoza

Georgia: Sophia Gorgodze, Tamar Bregvadze* and Natia Mzahavnadze*

Hong Kong (China): Ho-Pun Choi, Barry Lau, Fanny Yuen-Fan Wan* and Chun-Sing Woo*

Indonesia: Suprayitno Totok

Jordan: Abdalla Yousef Awad Al-Ababneh

Kazakhstan: Yerlikzhan Sabyruly, Serik Irsaliyev* and Nurgul Shamshieva*

Kosovo: Valmir Gashi

Lebanon: Nada Oweijane

Macao (China): Pak Sang Lou and Leong Lai*

Malaysia: Habibah Abdul Rahim, Dato Sri Khairil Awang* and Suliaman Wak*

Malta: Charles L. Carmelo Mifsud

Republic of Moldova: Anatolie Topala

Montenegro: Dragana Dmitrovic

Morocco: Mohammed Sassi

Republic of North Macedonia: Natasha Jankovska and Natasha Janevska*

Panama: Nadia De Leon and Marelisa Tribaldos*

Peru: Humberto Perez León Ibáñez and Liliana Miranda Molina*

Philippines: Nepomuceno A. Malaluan

Qatar: Khalid Abdulla Q. Al-Harqan

Romania: Daniela Bogdan*

Russian Federation: Sergey Kravtsov, Pavel Zenkovich and Anton Chetvertkov*

Saudi Arabia: Abdullah Alqataee, Husam Zaman, Nayyaf Al-Jabri, Mohamed Al-Harthi*, Faisal Mashary Al Saud* and Saja Jamjoom*

Serbia: Anamarija Vicek and Zorana Lužanin*

Singapore: Chern Wei Sng and Kwah Gek Low*

Chinese Taipei: Tian-Ming Sheu, Hwawei Ko* and Li-Chun Peng*

Ukraine: Sergiy Rakov, Inna Sovsun* and Pavlo Khobzey*

United Arab Emirates: Rabaa Alsumaiti, Hessa Alwahhabi, Ayesha Al Marri*, Khawla Al Mualla* and Moza Rashid Alghufli*

Uruguay: Andrés Peri

Viet Nam: Sai Cong Hong and My Ha Le Thi

PISA 2018 NATIONAL PROJECT MANAGERS

(*Former PISA 2018 NPM)

Albania: Rezana Vrapi

Argentina: Cecilia Beloqui and Samanta Bonelli*

Australia: Sue Thomson

Austria: Birgit Suchań

Azerbaijan: Narmina Aliyeva

Belarus: Jurij Miksiuk and Julia Khokhlova

Belgium: Inge De Meyer and Anne Matoul

Bosnia and Herzegovina: Žaneta Džumhur

Brazil: Aline Mara Fernandes

Brunei Darussalam: Hazri Kifle, Hjh Kamlah Hj Daud* and Habibah Hj Sion*

Bulgaria: Natalia Vassileva and Svetla Petrova*

Canada: Kathryn O'Grady, Tanya Scerbina and Pierre Brochu*

Chile: Ema Lagos Campos

Beijing-Shanghai-Jiangsu-Zhejiang (China): Tao Xin

Colombia: Natalia González Gómez and Andrés Gutiérrez Rojas*

Costa Rica: Rudy Masís Siles and Lilliam Mora Aguilar*

Croatia: Ana Markocic Dekanic and Michelle Bras Roth*

Czech Republic: Radek Blažek

Denmark: Hans Hummelgaard, Helga Foldbo, Vibeke Tornhøj Christensen and Óli Jákup Joensen*

Dominican Republic: Massiel Cohen Camacho

Estonia: Gunda Tire

Finland: Arto Ahonen

France: Irène Verlet

Georgia: Lasha Kokilashvili, Sophie Baxutashvili* and Tamar Bregvadze*

Germany: Kristina Reiss, Mirjam Weis and Christine Sälzer*

Greece: Ioannis Tsirmpas and Chryssa Sofianopoulou*

Hong Kong (China): Kit-Tai Hau

Hungary: László Ostorics

Iceland: Guðmundur Þorgrímsson, Almar Miðvik Halldórsson* and Svanhildur Steinarsdóttir*

Indonesia: Moch Abduh and Nizam Nizam*

Ireland: Caroline McKeown

Israel: Georgette Hilu, Inbal Ron-Kaplan and Joel Rapp*

Italy: Laura Palmerio

Japan: Yu Kameoka and Akiko Ono*

Jordan: Emad Ghassab Ababneh

Kazakhstan: Temirlan Kultumanov, Yerlikzhan Sabyruly, Magzhan Amangazy* and Irina Imanbek*

Korea: Seongmin Cho and Ku Jaok*

Kosovo: Mustafa Kadriu

Latvia: Andris Kangro

Lebanon: Bassem Issa

Lithuania: Natalija Valaviciene and Mindaugas Stundza*

Luxembourg: Bettina Boehm

Macao (China): Kwok-Cheung Cheung

Malaysia: Wan Raisuha Binti Wan Ali

Malta: Louis Scerri

Mexico: María Antonieta Díaz Gutierrez

Republic of Moldova: Valeriu Gutu and Anatolie Topala

Montenegro: Divna Paljevic

Morocco: Ahmed Chaibi

Netherlands: Joyce Gubbels, Martina Meelissen and Andrea Netten*

New Zealand: Adam Jang-Jones, Steven May and Saila Cowles*

Republic of North Macedonia: Beti Lameva

Norway: Fredrik Jensen and Marit Kjærnsli*

Panama: Ariel Melo, Jahir Calvo* and Genoveva Iglesias*

Peru: Humberto Perez León Ibáñez and Liliana Miranda*

Philippines: Nelia Vargas Benito

Poland: Barbara Ostrowska

Portugal: Vanda Lourenço* and João Maroco Domingos*

Qatar: Shaikha Al-Ishaq

Romania: Simona Velea

Russian Federation: Galina Kovaleva

Saudi Arabia: Fahad Abdullah Alharbi and Mohammed Al-Sobeiy*

Serbia: Gordana Capric and Dragica Pavlovic-Babic*

Singapore: Elaine Chua and Chew Leng Poon*

Slovak Republic: Julia Miklovicova and Jana Ferencová*

Slovenia: Klaudija Šterman Ivančič and Mojca Štraus*

Spain: Lis Cercadillo

Sweden: Ellen Almgren, Eva Lundgren* and Agnes Tongur*

Switzerland: Andrea B. Erzinger and Christian Nidegger*

Chinese Taipei: Pi-Hsia Hung

Thailand: Ekarin Achakunwisut

Turkey: Umut Erkin Taş

Ukraine: Tetiana Vakulenko and Anna Novosad*

United Arab Emirates: Shaikha Al Zaabi, Ahmed Hosseini and Moza Rashid Al Ghufli

United Kingdom: Juliet Sizmur

United States: Patrick Gonzales

Uruguay: María Helvecia Sánchez Núñez

Viet Nam: My Ha Le Thi

OECD SECRETARIAT

Andreas Schleicher (Strategic development)

Marilyn Achiron (Editorial support)

Alejandra Arbeláez Ayala (Analytic services)

Francesco Avvisati (Analytic services)

Yuri Belfali (Strategic development)

Simone Bloem (Dissemination support)

Guillaume Bousquet (Analytic services)

Alison Burke (Production support)

Cassandra Davis (Dissemination co-ordination)

Alfonso Echazarra (Analytic services)

Juliet Evans (Communication & dissemination)

Natalie Foster (Analytic services)

Pauline Givord (Analytic services)

Hélène Guillou (Analytic services)

Tue Halgreen (Project management)

Parker Hart (Dissemination support)

Julia Himstedt (Communication & dissemination)

Miyako Ikeda (Analytic services)

Natalie Laechelt (Project management)

Sophie Limoges (Production support)

Camille Marec (Analytic services)

Thomas Marwood (Administrative support)

Nicolás Miranda (Analytic services)

Jeffrey Mo (Analytic services)

Chiara Monticone (Analytic services)

Tarek Mostafa (Analytic services)

Tomoya Okubo (Analytic services)

Lesley O'Sullivan (Administrative support)

Judit Pál (Analytic services)

Mario Piacentini (Analytic services)

Giannina Rech (Analytic services)

Daniel Salinas (Analytic services)

Markus Schwabe (Analytic services)

Della Shin (Production support)

Rebecca Tessier (Production support)

Hanna Varkki (Administrative support)

Sophie Vayssettes (Project management)

PISA 2018 READING EXPERT GROUP

Core Expert Group

Jean-François Rouet (Chair) (University of Poitiers, France)

Paul van den Broek (Leiden University, The Netherlands)

Kevin Kien Hoa Chung (The Education University of Hong Kong China)

Dominique Lafontaine (QEG Liaison) (University of Liège, Belgium)

John Sabatini (Educational Testing Service, United States)

Sascha Schroeder (University of Cologne, Germany)

Sari Sulkunen (University of Jyväskylä, Finland)

Extended Expert Group

Gina Biancarosa (University of Oregon, United States)

Ivar Braten (University of Oslo, Sweden)

Marina I. Kuznetkova (Russian Academy of Education, Russia)

Nele McElvany (Technische Universität Dortmund, Germany)

Eduardo Vidal-Abarca (University of Valencia, Spain)

William G. Brozo (University of South Carolina, United States)

Kate Cain (Lancaster University, United Kingdom)

PISA 2018 GLOBAL COMPETENCE EXPERT GROUP

Experts who led the first phase of development

David Kerr (University of Reading and Young Citizens, United Kingdom)

Peter Franklin (HTWG Konstanz University of Applied Sciences, Germany)

Darla Deardorff (Duke University, United States)

Sarah Howie (University of Stellenbosch, South Africa)

Wing On Lee (Open University of Hong Kong, China)

Jasmine B.-Y. Sim (National Institute of Education, Singapore)

Sari Sulkunen (Jyväskylä University, Finland)

Experts who led the second phase of development

Martyn Barrett (Chair) (University of Surrey, United Kingdom)

Veronica Boix Mansilla (Harvard University, United States)

Darla Deardorff (Duke University, United States)

Hye-Won Lee
(Korea Institute for Curriculum and Evaluation [KICE], Korea)

Extended group

Tom Franklin (Young Citizens, United Kingdom)

Alicia Cabezudo (Universidad Nacional de Rosario, Argentina)

Hans Ruesink (Ministry of Education, Culture and Science, The Netherlands)

Myunghee Ju Kang (Ewha Womans University, South Korea)

Jom Schreiber (Duquesne University, United States)

Jo-Anne Baird (University of Oxford, United Kingdom)

Naomi Miyake (University of Tokyo, Japan)

PISA 2018 QUESTIONNAIRE EXPERT GROUP

Core Expert Group

Fons J. R. van de Vijver (Chair) (Tilburg University, the North-West University and the University of Queensland, The Netherlands and Australia)

Dominique Lafontaine (University of Liège, Belgium)

David Kaplan (University of Wisconsin, United States)

Sarah Howie (University of Stellenbosch, South Africa)

Andrew Elliot (University of Rochester, United States)

Therese Hopfenbeck (Oxford University, England)

Extended Expert Group

David Cantor (University of London, United Kingdom)

Kit-Tai Hau (The Chinese University of Hong Kong, China)

Hwa-Wei Ko (National Central University, Chinese Taipei)

Malgorzata Mikucka (Universität Mannheim, Germany)

Naomi Miyake (University of Tokyo, Japan)

Thierry Rocher (Ministère de l'Éducation Nationale, France)

Herb Marsh (Australian Catholic University, Australia)

Ben Jensen (Learning First, Australia)

Technical Advisory Group

Keith Rust (chair) (Westat, United States)

Kentaro Yamamoto (ETS, United States)

John de Jong (VU University Amsterdam, Netherlands)

Christian Monseur (University of Liège, Belgium)

Leslie Rutkowski (University of Oslo, Norway and Indiana University, United States)

Cees Glas (University of Twente, Netherlands)

Irwin Kirsch (ETS, United States)

Theo Eggen (Cito, Netherlands)

Kit-Tai Hau (The Chinese University of Hong Kong, China)

Oliver Lüdtke (IPN - Leibniz Institute for Science and Mathematics Education, Germany)

Matthias von Davier (NBME, United States)

David Kaplan (University of Wisconsin – Madison, United States)

Thierry Rocher (Ministère de l'Éducation Nationale, France)

Margaret Wu (Victoria University, Australia)

PISA 2018 LEAD CONTRACTORS

Educational Testing Service (United States) – Core A lead contractor

Irwin Kirsch (International Project Director)

Claudia Tamassia (International Project Manager)

David Garber (Project Management)

Ann Kennedy (Project Management)

Larry Hanover (Editorial Support)

Lisa Hemat (Project Support)

Isabelle Jars (Project Management, Questionnaires)

Luisa Langan (Project Management, Questionnaires)

Judy Mendez (Project Support and Contracts)

Daniel Nicastro (Project Support)

Yelena Shuster (Project Support)

Eugenio Gonzalez (Training and Data Poducts)

Kentaro Yamamoto (Director, Psychometrics and Analysis)

Fred Robin (Manager, Psychometrics and Analysis)

Usama Ali (Psychometrics and Analysis)

Haiwen Chen (Psychometrics and Analysis)

Qiwei He (Psychometrics and Analysis)

Sean-Hwane Joo (Psychometrics and Analysis)

Lale Khorramdel (Psychometrics and Analysis)

Selene Sunmin Lee (Psychometrics and Analysis)

Emily Lubaway (Psychometrics and Analysis)

Hyo Jeong Shin (Psychometrics and Analysis)

Peter van Rijn (Psychometrics and Analysis)

Laura Halderman (Lead Test Developer and Test Development Coordinator, Reading Literacy and Global Competence)

Kelly Bruce (Test Developer and Test Development Coordinator, Reading Literacy)

Marylou Lennon (Test Developer and Test Development Coordinator, Global Competence)

Patti Mendoza (Test Developer, Reading Literacy)

Eric Miller (Test Developer, Reading Literacy)

Laura Shook (Test Developer, Reading Literacy)

Denise Walker (Test Developer, Reading Literacy)

James Seal (Test Developer, Reading Literacy)

Darla Scates (Test Developer, Reading Literacy)

Scott Seay (Test Developer, Reading Literacy)

John Fischer (Test Developer, Reading Literacy)

Nial Eastman (Reviewer, Reading Litearcy)

Mary Kathryn Arnold (Reviewer, Reading Literacy)

Lynette Perloff (Reviewer, Reading Literacy)

John Hawthron (Test Developer, Global Competence)

Douglas Baldwin (Test Developer, Global Competence)

Tenaha O'Reilly (Test Developer, Global Competence)

Michael Wagner (Director, Platform Development)

Jason Bonthron (Platform Development and Authoring)

Paul Brost (Platform Development)

Ramin Hemat (Platform Development and Authoring)

Keith Keiser (Platform Development and Coding System)

Debbie Pisacreta (Interface Design and Graphics)

Janet Stumper (Graphics)

Chia Chen Tsai (Platform Development)

Ted Blew (Area Director, Data Analysis and Research Technologies)

John Barone (Director, Data Analysis and Database Technologies)

Mathew Kandathil (Team Leader, Data Analysis and Data Management)

Kevin Bentley (Data Products)

Hezekiah Bunde (Data Management)

Karen Castellano (Data Analysis)

Matthew Duchnowski (Data Management)

Ying Feng (Data Management)

Harrison Gamble (Data Analysis)

Zhumei Guo (Data Analysis)

Paul Hilliard (Data Analysis)

Lokesh Kapur (Data Analysis)

Debra Kline (Project Management)

Phillip Leung (Data Quality, Data Products)

Alfred Rogers (Data Management, Data Products)

Carla Tarsitano (Project Management)

Tao Wang (Data Quality)

Lingjun Wong (Data Analysis)

Ping Zhai (Data Analysis)

Wei Zhao (Data Analysis)

Pearson (United Kingdom) – Core B lead contractor

John de Jong (Programme Director)

Peter Foltz (Content lead, Reading Literacy)

Christine Rozunick (Content lead, Background Questionnaire)

Jon Twing (Psychometric consultant)

Dave Leach (Programme Manager and Programme Director)

Lorraine Greenwood (Project management)

Jay Larkin (Editor and support for Reading literacy)

Madison Cooper (Editor and support for Background Questionnaire)

Clara Molina (Programme Administrator)

Mark Robeck (Minutes and editor)

Kimberly O'Malley (Additional management support)

Westat (United States) – Core C lead contractor

Keith Rust (Director of the PISA Consortium for Sampling and Weighting)

Sheila Krawchuk (Sampling and Weighting)

Jessica Chan (Sampling)

David Ferraro (Weighting)

Susan Fuss (Sampling and Weighting)

Moriah Goodnow (Weighting)

Amita Gopinath (Weighting)

Jing Kang (Sampling and Weighting)

Véronique Lieber (Sampling and Weighting)

John Lopdell (Sampling and Weighting)

Neha Patel (Weighting)

Shawn Lu (Weighting)

Jacqueline Severynse (Sampling and Weighting)

Yumiko Siegfried (Sampling and Weighting)

Joel Wakesberg (Sampling and Weighting)

Sipeng Wang (Sampling)

Natalia Weil (Sampling and Weighting)

Erin Wiley (sampling and Weighting)

Sergey Yagodin (Weighting)

cApStAn Linguistic Quality Control (Belgium) – Core D lead contractor

Steve Dept (Project Director, Translatability Assessment)

Lieve Deckx (Verification Management, Cognitive Units)

Andrea Ferrari (Linguistic Quality Assurance and Quality Control Designs)

Musb Hayatli (Right-to-Left Scripts, Cultural Adaptations)

Emel Ince (Verification Management, Manuals)

Elica Krajceva (Verification Management, Questionnaires)

Shinoh Lee (Verification Management, Cognitive Units)

Irene Liberati (Verification Management, Cognitive Units)

Roberta Lizzi (Verification Management, Trend Content)

Manuel Souto Pico (Translation Technologist, Linguistic Quality Assurance Tools and Procedures)

Laura Wayrynen (Lead Project Manager)

PISA 2018 CONTRIBUTORS, WORKING WITH LEAD CONTRACTORS

Australian Council for Educational Research (Australia) – Core C contributor

Eveline Gebhardt (Project Director)

Bethany Davies (School Sampling)

Jorge Fallas (School and Student Sampling)

Jennifer Hong (School Sampling)

Renee Kwong (School and Student Sampling)

Dulce Lay (School Sampling)

Gregory Macaskill (School Sampling)

Martin Murphy (School Sampling)

Claire Ozolins (School Sampling)

Leigh Patterson (School Sampling)

Alla Routitsky (Student Sampling)

BranTra (Belgium) – Core D contributor

Eva Jacob (Translation Management, French Source Development)

Danina Lupsa (Translation Technologist, Linguistic Quality Assurance Tools and Procedures)

Ben Meessen (Translation Management, Development of Common Reference Versions for Spanish, Chinese, Arabic)

Deutsches Institut für Internationale Pädagogische Forschung (DIPF, Germany – Core A contributor on the development of the questionnaires)

Eckhard Klieme (Study Director, Questionnaire Framework and Development)

Nina Jude (Management and Questionnaire Development)

Sonja Bayer (Questionnaire Development and Analysis)

Janine Buchholz (Questionnaire Scaling)

Frank Goldhammer (Questionnaire Development)

Silke Hertel (Questionnaire Development)

Franz Klingebiel (Questionnaire Development)

Susanne Kuger (Questionnaire Framework and Development)

Ingrid Mader (Team Assistance)

Tamara Marksteiner (Questionnaire Analysis)

Jean-Paul Reeff (International Consultant)

Nina Roczen (Questionnaire Development)

Brigitte Steinert (Questionnaire Development)

Svenja Vieluf (Questionnaire Development)

HallStat SPRL (Belgium) – Core A contributor as the translation referee

Béatrice Halleux (Consultant, Translation/Verification Referee, French Source Development)

Statistics Canada (Canada) – Core A DIPF contributor on questionnaires

Sylvie Grenier (Overall Management)

Patrick Cloutier (Implementation Delivery System)

Ginette Grégoire (Implementation Delivery System)

Martine Lafrenière (Implementation Delivery System)

Rosa Tatasciore (Implementation Delivery System)

Unité d'analyse des Systèmes et des Pratiques d'enseignement (aSPe, Belgium) – Core A contributor on coding training

Dominique Lafontaine (Project Supervisor)

Anne Matoul (Coding Training, Reading)

Stéphanie Géron (Coding Training, Reading)

Valérie Bluge (Coding Training, Reading)

Valérie Quittre (Coding Training, Science)

Isabelle Demonty (Coding Training, Mathematics)

University of Luxembourg (Luxembourg) – Core A contributor on test development

Romain Martin (Test Development Coordinator)

Samuel Greiff (Test Development Coordinator)

Antoine Fischbach (Test Development Coordinator)

Robert Reuter (Test Development)

Monique Reichert (Test Development)

Philipp Sonnleitner (Test Development)

Christoph Kemper (Test Development)

Maida Mustafic (Test Development)

Purya Baghaei (Test Development)

Vincent Koenig (User Testing)

Sophie Doublet (User Testing)

Westat (United States) – Core A contributor on survey operations

Merl Robinson (Director of Core A Contractor for Survey Operations)

Michael Lemay (Manager of Core A Contractor for Survey Operations)

Sarah Sparks (National Centre Support, Quality Control)

Beverley McGaughan (National Centre Support, Quality Control)

PISA 2018 Results (Volume I)
WHAT STUDENTS KNOW AND CAN DO

The OECD Programme for International Student Assessment (PISA) examines what students know in reading, mathematics and science, and what they can do with what they know. It provides the most comprehensive and rigorous international assessment of student learning outcomes to date. Results from PISA indicate the quality and equity of learning outcomes attained around the world, and allow educators and policy makers to learn from the policies and practices applied in other countries. This is one of six volumes that present the results of the PISA 2018 survey, the seventh round of the triennial assessment.

Volume I, *What Students Know and Can Do,* provides a detailed examination of student performance in reading, mathematics and science, and describes how performance has changed since previous PISA assessments.

Volume II, *Where All Students Can Succeed,* examines gender differences in student performance, and the links between students' socio-economic status and immigrant background, on the one hand, and student performance and well-being, on the other.

Volume III, *What School Life Means for Students' Lives,* focuses on the physical and emotional health of students, the role of teachers and parents in shaping the school climate, and the social life at school. The volume also examines indicators of student well-being, and how these are related to the school climate.

Volume IV, *Are Students Smart about Money?,* examines 15-year-old students' understanding about money matters in the 21 countries and economies that participated in this optional assessment.

Volume V, *Effective Policies, Successful Schools,* analyses the policies and practices used in schools and school systems, and their relationship with education outcomes more generally.

Volume VI, *Are Students Ready to Thrive in Global Societies?,* explores students' ability to examine local, global and intercultural issues, understand and appreciate different perspectives and world views, interact respectfully with others, and take responsible action towards sustainability and collective well-being.

Consult this publication on line at: ***https://doi.org/10.1787/5f07c754-en***

This work is published on the *OECD iLibrary,* which gathers all OECD books, periodicals and statistical databases.
Visit ***www.oecd-ilibrary.org*** for more information.

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-46038-6

9 789264 460386