



## **Assessments, evaluations and how they are used**

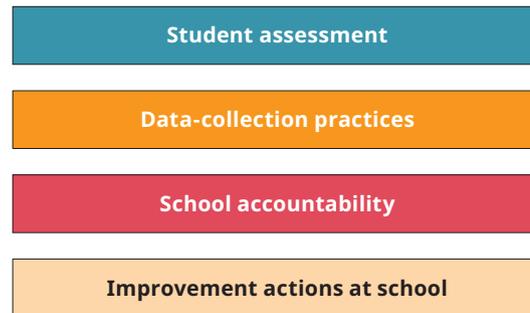
This chapter examines four evaluation and assessment activities in particular: student assessments, data-collection practices, school accountability and improvement actions at school. It discusses how school systems use the information they gather from these evaluations and assessments. The ways school systems use this data are then correlated with student performance and equity in the education system.

Evaluation and assessment, as discussed in this chapter, refers to the policies and practices through which education systems assess student learning, and evaluate teacher practices and school outcomes. It also conveys how education systems and schools use the results from assessment and evaluation to improve classroom processes and student learning.

The *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes* identifies the hallmarks of strong evaluation and assessment systems (OECD, 2013<sub>[1]</sub>). These include setting clear and ambitious goals or standards for what is expected of students, schools and the system overall, and collecting reliable data to measure the extent to which goals are being met. In addition, in strong evaluation and assessment systems, students, teachers, schools and policy makers receive the feedback they need to reflect critically on their own progress, and remain engaged and motivated to succeed.

In this chapter, four evaluation and assessment topics are covered (Figure V.8.1): student assessment, data-collection practices, school accountability and improvement actions at school.

Figure V.8.1 **Evaluation and assessment as covered in PISA 2018**



## What the data tell us

- On average across OECD countries, the prevalence of using student assessments for various purposes declined between 2012 and 2018. For example, in 2012, 55% of students were enrolled in a school that compares its performance with that of other schools, while in 2018, 46% of students attended such a school. Similarly, in 2012, 53% of students attended a school that uses student assessments to make judgements about teachers' effectiveness, while in 2018, 44% of students attended such a school.
- Some 38% of students were enrolled in schools that post achievement data publicly, on average across OECD countries. These students scored five points higher in reading, on average across OECD countries, than students in schools that do not post data publicly, even after accounting for students' and schools' socio-economic profile. At the system level, however, the incidence of posting achievement data publicly was not correlated with mean performance in PISA, nor with equity in performance.
- On average across OECD countries, students in schools whose principal reported that their school seeks written feedback from students scored better in reading than students in schools that do not seek written feedback, even after accounting for students' and schools' socio-economic profile. In addition, equity in student performance tended to be greater amongst countries and economies that have a higher percentage of students in schools that seek written feedback from students.
- High-performing countries and economies tended to have more teacher mentoring on the school's initiative. In those systems, more schools implemented a standardised policy for reading-related subjects taught at school (including a school curriculum with shared instructional materials, and staff development and training) based on district or national policies.
- Countries and economies tended to have better equity in education when they: use student assessments to inform parents about their child's progress; use student assessments to identify aspects of instruction or the curriculum that could be improved; use written specifications for student performance on the school's initiative; seek feedback from students; and have regular consultations on school improvement at least every six months, based on district or national policies.

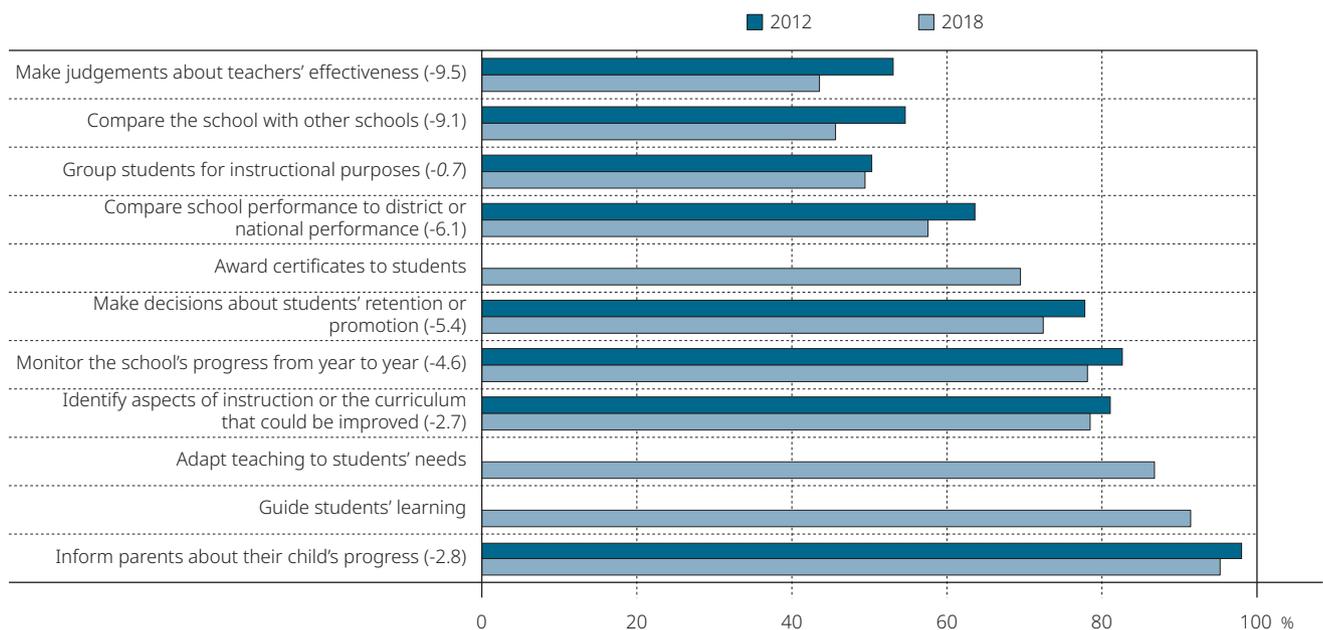
## STUDENT ASSESSMENT

Student assessment refers to judgements on individual students' progress and achievement. It covers school-based assessments designed by teachers or other staff at school, as well as large-scale external assessment and examinations (OECD, 2013<sup>[1]</sup>). Student assessments can have different purposes (Rosenkvist, 2010<sup>[2]</sup>; Lesaux, 2006<sup>[3]</sup>; Looney, 2011<sup>[4]</sup>). For example, formative assessments, or assessments for learning, do not have consequences for students, but instead aim to provide feedback to help students progress on their learning path (Shepard, 2000<sup>[5]</sup>; Hattie and Timperley, 2007<sup>[6]</sup>). Summative assessments, or assessments of learning, summarise and certify achieved learning and are sometimes used to make high-stakes decisions for students or teachers, such as promoting or retaining students, or grouping students by their achievement level.

PISA 2018 asked school principals about the purposes of the student assessments used in their school. On average across OECD countries, student assessments were most often used to inform parents about their child's progress (95% of students were in schools whose principal reported that assessments are used for this purpose), to guide students' learning (91%), and to adapt teaching to the students' needs (87%) (Figure V.8.2). Student assessments were also commonly used to identify aspects of instruction or the curriculum that could be improved (78%), to monitor the school's progress from year to year (78%) and to make decisions about students' retention or promotion (72%). Less than half of students were in schools that use student assessments to group students by ability, to compare the school to other schools, or to make judgements about teachers' effectiveness, on average across OECD countries.

Figure V.8.2 **Change between 2012 and 2018 in purposes of student assessment**

*Percentage of students in schools whose principal reported that assessments of students are used for the following; OECD average*



**Notes:** Differences between 2012 and 2018 are shown between parentheses next to the purpose of student assessment. All differences are statistically significant, except for the item "Group students for instructional purposes" (see Annex A3). Some items do not have a dark blue bar because data were not collected in PISA 2012.

**Source:** OECD, PISA 2018 Database, Table V.B1.8.1.

**StatLink**  <https://doi.org/10.1787/888934131899>

According to school principals, using student assessments for formative purposes was common in nearly all PISA-participating countries and economies. Three items included in Figure V.8.3 are measures of formative student assessment:

- In 2018, at least 70% of students in every participating country and economy were enrolled in schools that use student assessments to guide student learning (Table V.B1.8.1).
- Similarly, in 68 out of 79 countries and economies at least 70% of students were enrolled in schools that use student assessments to identify aspects of instruction or the curriculum that could be improved.
- In addition, in 74 countries and economies, at least 70% of students attended a school that uses student assessments to adapt instruction to students' needs.

By contrast, the incidence of using student assessments to make high-stakes decisions for either students or teachers was somewhat lower:

- In only 22 countries and economies were there at least 70% of students enrolled in schools that use student assessments to group students for instructional purposes.
- In 36 countries and economies, at least 70% of students were enrolled in schools that use student assessments to make judgements about teachers' effectiveness.
- In 58 countries and economies, at least 70% of students were enrolled in schools that use student assessments to make decisions about retaining or promoting students.

On average across OECD countries, the practice of using student assessments to make judgements about teachers' effectiveness declined by ten percentage points between 2012 and 2018, using student assessments to compare the school with other schools declined by nine percentage points, and using student assessments to make decisions about students' retention or promotion declined by five percentage points (Table V.B1.8.1). Of all the purposes of student assessments included in Figure V.8.2, using student assessments for grouping students for instructional purposes – a practice that remained stable over the period (50% of students attended schools that use student assessments for this purpose) – was the only practice that did not decline over the period, on average across OECD countries.

The prevalence of using student assessments for specific purposes was similar between socio-economically advantaged and disadvantaged schools, with some exceptions. On average across OECD countries, using student assessments to group students for instructional purposes, to adapt teaching to students' needs and to monitor the school's progress from year to year was more common amongst disadvantaged than advantaged schools (Table V.B1.8.2).

PISA measures of the purposes of student assessments included in Figure V.8.2 were unrelated or only weakly related to reading performance (Table V.B1.8.4). On average across OECD countries, schools that showed higher performance tended to use student assessments to make decisions about their students' retention or promotion, after accounting for students' and schools' socio-economic profile. In contrast, schools that showed lower performance tended to use student assessments to guide their students' learning or adapt teaching to their students' needs, on average across OECD countries. However, these results were not consistent across countries, showing both positive and negative associations.

## COLLECTING DATA ON STUDENT OUTCOMES

PISA 2018 asked school principals whether their school collected data on student outcomes, specifically on students' test results and graduation rates. On average across OECD countries, some 93% of students were in schools whose principals reported that they systematically record students' test results and graduation rates (Figure V.8.3). Only in Austria, Brazil, Finland, Greece, Luxembourg and Switzerland was the share of students in schools that carried out systematic recording of student outcomes lower than 85%.

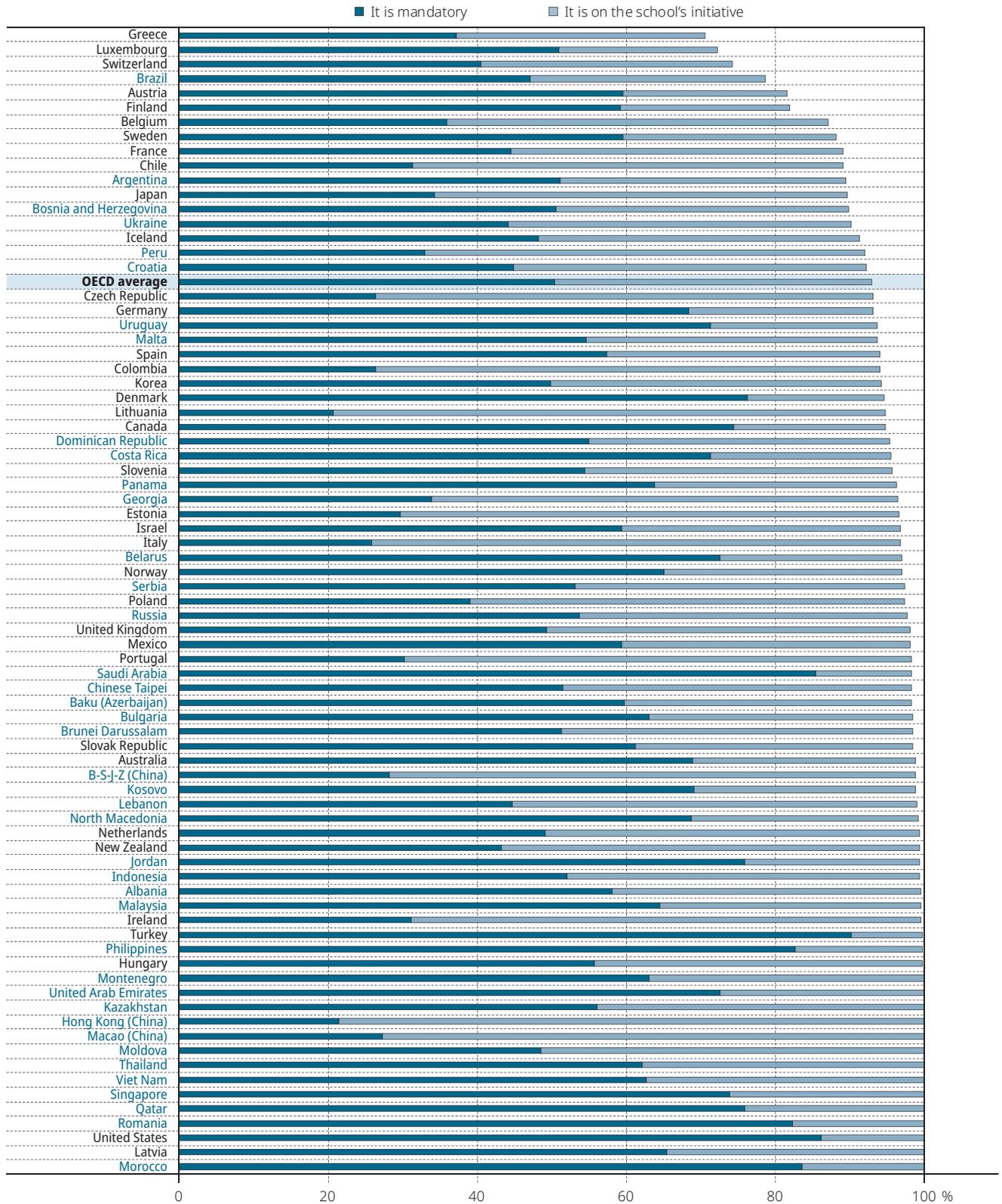
On average across OECD countries, about half of students were in schools whose principal reported that collecting data on student outcomes is mandatory (e.g. based on district or ministry policies), and 42% were in schools where data is collected on the school's initiative (Figure V.8.2). In most countries and economies, most students were in schools where the collection of data on student outcomes is mandatory; but in 20 countries and economies, more than half of students were in schools where data on student outcomes is collected on the school's initiative. In Beijing, Shanghai, Jiangsu and Zhejiang (China) (hereafter "B-S-J-Z [China]"), Hong Kong (China), Italy, Lithuania and Macao (China), more than 70% of students were in schools where data on student outcomes is collected on the school's initiative, as opposed to being based on mandatory governmental policies.

The systematic recording of student outcomes was more prevalent in 2018 than in 2015 in Italy, Kosovo, Lithuania, Montenegro, the Republic of North Macedonia (hereafter "North Macedonia") and Switzerland, and was less prevalent in Brazil, Greece, Iceland, Luxembourg, Malta and Slovenia (Table V.B1.8.12).

On average across OECD countries, the prevalence of collecting data on student outcomes was similar, regardless of the socio-economic profile of the school, but differences were observed in some countries. In Austria, Belarus, Brunei Darussalam, Georgia, Iceland, Peru and Spain, collecting data on students' results and graduation rates was more prevalent in advantaged schools, while in Denmark, France, Kazakhstan, Lithuania, Luxembourg and North Macedonia, this practice was more prevalent in disadvantaged schools (Table V.B1.8.13).

Figure V.8.3 **Collecting data on student outcomes in schools**

Percentage of students in schools that systematically record students' test results and graduation rates; based on principals' reports



Countries are ranked in ascending order of the percentage of students in schools where such data are collected.

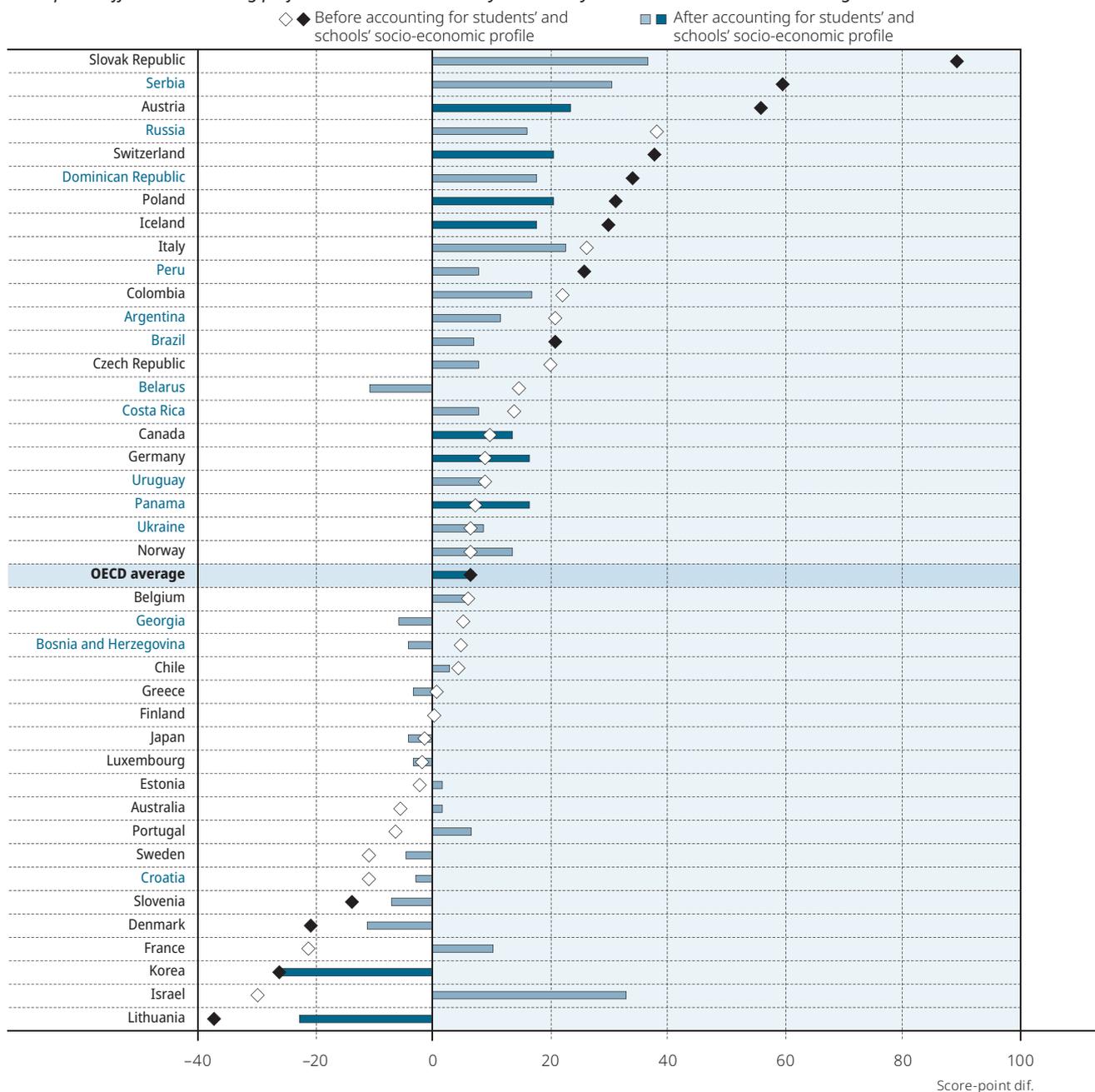
Source: OECD, PISA 2018 Database, Table V.B1.8.11.

StatLink <https://doi.org/10.1787/888934131918>

On average across OECD countries, students in schools that systematically record students' test results and graduation rates scored better in reading than students in schools that do not collect this kind of data (a difference of six score points). After accounting for students' and schools' socio-economic profile, students in schools that collect data on student outcomes scored better in reading, on average across OECD countries (a difference of seven points) and in seven countries and economies (Figure V.8.4).

Figure V.8.4 **Recording students' test results and graduation rates, and reading performance**

Score-point difference in reading performance in schools that systematically record students' test results and graduation rates



**Notes:** Statistically significant changes are shown in darker tones (see Annex A3).

The socio-economic profile is measured by the PISA index of economic, social and cultural status.

Countries and economies are ranked in descending order of the score-point difference in reading performance in schools that systematically record students' test results and graduation rates, before accounting for students' and schools' socio-economic profile.

**Source:** OECD, PISA 2018 Database, Table V.B1.8.15.

**StatLink** <https://doi.org/10.1787/888934131937>

## SCHOOL ACCOUNTABILITY

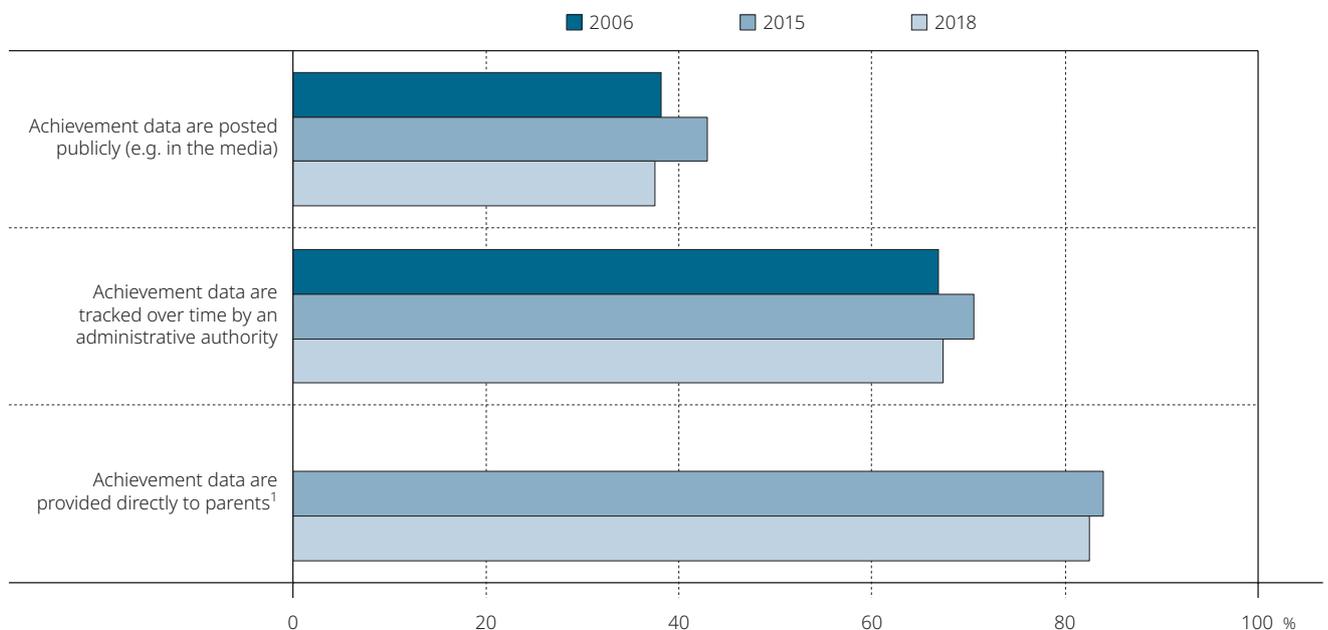
PISA 2018 asked school principals whether their school uses achievement data for accountability purposes, and if so, how. Three forms of school accountability were considered: providing achievement data to parents directly, posting achievement data publicly (for example, in the media), and tracking achievement data over time by an administrative authority. Achievement data, as defined here, refers to aggregated school or grade-level test scores or grades, or graduation rates.

On average across OECD countries in 2018, about 83% of students attended schools that provide achievement data to parents directly (Figure V.8.5). This is the most common form of school accountability considered in PISA. In all countries and economies except Austria, at least half of students were enrolled in schools that provide achievement data to parents directly (Table V.B1.8.7).

Posting achievement data publicly (for example, in the media) is the least common form of school accountability, on average across OECD countries (38% of students attended schools that do this) (Figure V.8.6). In 58 out of 79 education systems, less than half of students were in schools that post achievement data publicly (Table V.B1.8.7). Posting data publicly was more common amongst advantaged schools (43% of students attended such schools) than disadvantaged schools (34%), and more common in urban (39%) than in rural (33%) schools, on average across OECD countries (Table V.B1.8.8).

Figure V.8.5 Trends in school accountability, 2006, 2015 and 2018

Percentage of students in schools that use achievement data for accountability purposes; OECD average



1. Data on "Achievement data are provided directly to parents" were not collected in PISA 2006.

**Note:** All differences between 2015 and 2018 are statistically significant; all differences between 2006 and 2018 are not (see Annex A3).

**Source:** OECD, PISA 2018 Database, Table V.B1.8.7.

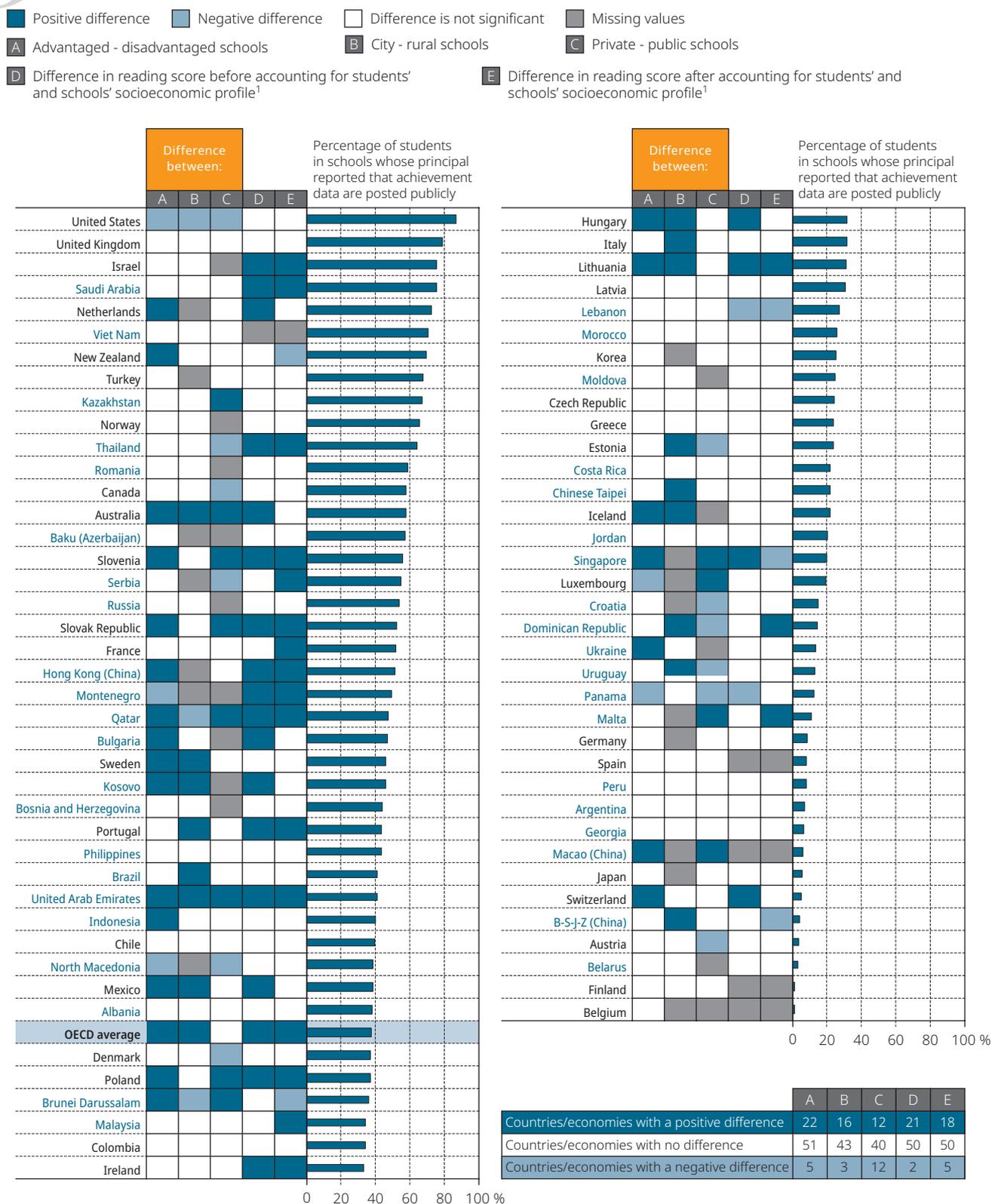
**StatLink**  <https://doi.org/10.1787/888934131956>

The earliest cycle in which PISA collected information on school accountability was PISA 2006. On average across OECD countries, there was no change between 2006 and 2018 in the share of students in schools that post achievement data publicly (Figure V.8.6). In Brazil, Bulgaria, Indonesia, Ireland, Israel, Norway, Portugal, the Slovak Republic, Slovenia and Turkey, posting achievement data was more prevalent in 2018 than in 2006; but in Croatia, the Czech Republic, Estonia, Luxembourg, Montenegro, the Netherlands, the Russian Federation, Sweden, Chinese Taipei and the United Kingdom, it was less prevalent in 2018 than in 2006 (Table V.B1.8.7).

Similarly, the share of students in schools where achievement data are tracked over time by an administrative authority did not change between 2006 and 2018, on average across OECD countries (in both cycles it was 67%) (Figure V.8.5). In 19 countries and economies, the share of students in schools where achievement data are tracked over time by an administrative authority increased (Table V.B1.8.7). The largest increases were observed in Denmark, Indonesia, Norway and Chinese Taipei. In 10 countries and economies, the share of students in schools where achievement data are tracked over time by an administrative authority shrank over the period. The largest decreases were observed in Finland, Iceland and Luxembourg.

PISA 2006 did not collect information about schools providing data to parents.

Figure V.8.6 Posting achievement data publicly, school characteristics and reading performance



1. This analysis is restricted to schools with the modal ISCED level for 15-year-old students.

Countries and economies are ranked in descending order of the percentage of students in schools whose principal reported that achievement data are posted publicly.

Source: OECD, PISA 2018 Database, Table V.B1.8.8.

StatLink <https://doi.org/10.1787/888934131975>

PISA also collected data on school accountability in 2015. All three forms of school accountability were less prevalent in 2018 than in 2015, on average across OECD countries. The share of students in schools that post achievement data publicly was five percentage points smaller in 2018 than in 2015, on average across OECD countries. In Argentina, Bulgaria, Croatia, Greece, Hong Kong (China), Iceland and Latvia, the share of students enrolled in schools that post data publicly decreased by ten percentage points or more; in Korea and Luxembourg, the shares shrank by 20 and 21 percentage points, respectively (Table V.B1.8.7). By contrast, in Kosovo, Macao (China), Malta, Montenegro, Slovenia and Chinese Taipei, the share of students in schools that post achievement data publicly increased by five percentage points or more during this period.

A majority of students were in schools that use at least one of the three forms of school accountability: providing achievement data to parents directly, posting achievement data publicly, and having achievement data tracked over time by an administrative authority (Table V.B1.8.6). Across OECD countries, on average, only 5% of students were in schools whose principal reported that none of these three forms is used. While less than 5% of students in more than 60 countries and economies attended such schools, more than 20% of students in Austria, Finland and Germany did. In contrast, 26% of students were in schools whose principal reported that all three forms of school accountability are used. More than 50% of students in Kazakhstan, New Zealand, Saudi Arabia, Thailand, Turkey, the United Kingdom, the United States and Viet Nam attended such schools.

On average across OECD countries, students enrolled in schools that post achievement data publicly scored better in reading (by 13 points) than students in schools that do not post data publicly. Yet, socio-economically advantaged schools were more likely than disadvantaged schools to post achievement data publicly. After accounting for students' and schools' socio-economic profile, students in schools that post achievement data publicly still scored five points higher in reading, on average across OECD countries (Figure V.8.6). At the system level, however, the incidence of posting achievement data publicly was not correlated with mean performance in PISA, nor with equity in performance, as described in detail later in this chapter.

The relationship between student achievement and other forms of school accountability was weak after accounting for students' and schools' socio-economic profile. After accounting for students' and schools' socio-economic profile, in nine countries and economies, students in schools where achievement data is tracked over time by an administrative authority scored better in reading; in two countries, they scored worse (Table V.B1.8.9).

Similarly, after accounting for students' and schools' socio-economic profile, in six countries, students in schools that provide achievement data to parents directly scored better in reading, but in three countries they scored worse (Table V.B1.8.10).

## IMPROVEMENT ACTIONS AT SCHOOL

### Seeking written feedback from students

One way that schools evaluate themselves is by seeking written feedback from students. On average across OECD countries, 68% of students were in schools whose principal reported that the school seeks written feedback from students regarding their lessons, teachers or resources. This practice is typically based on the school's own initiative (OECD average = 56%) rather than being mandatory (12%) (Table V.B1.8.11).

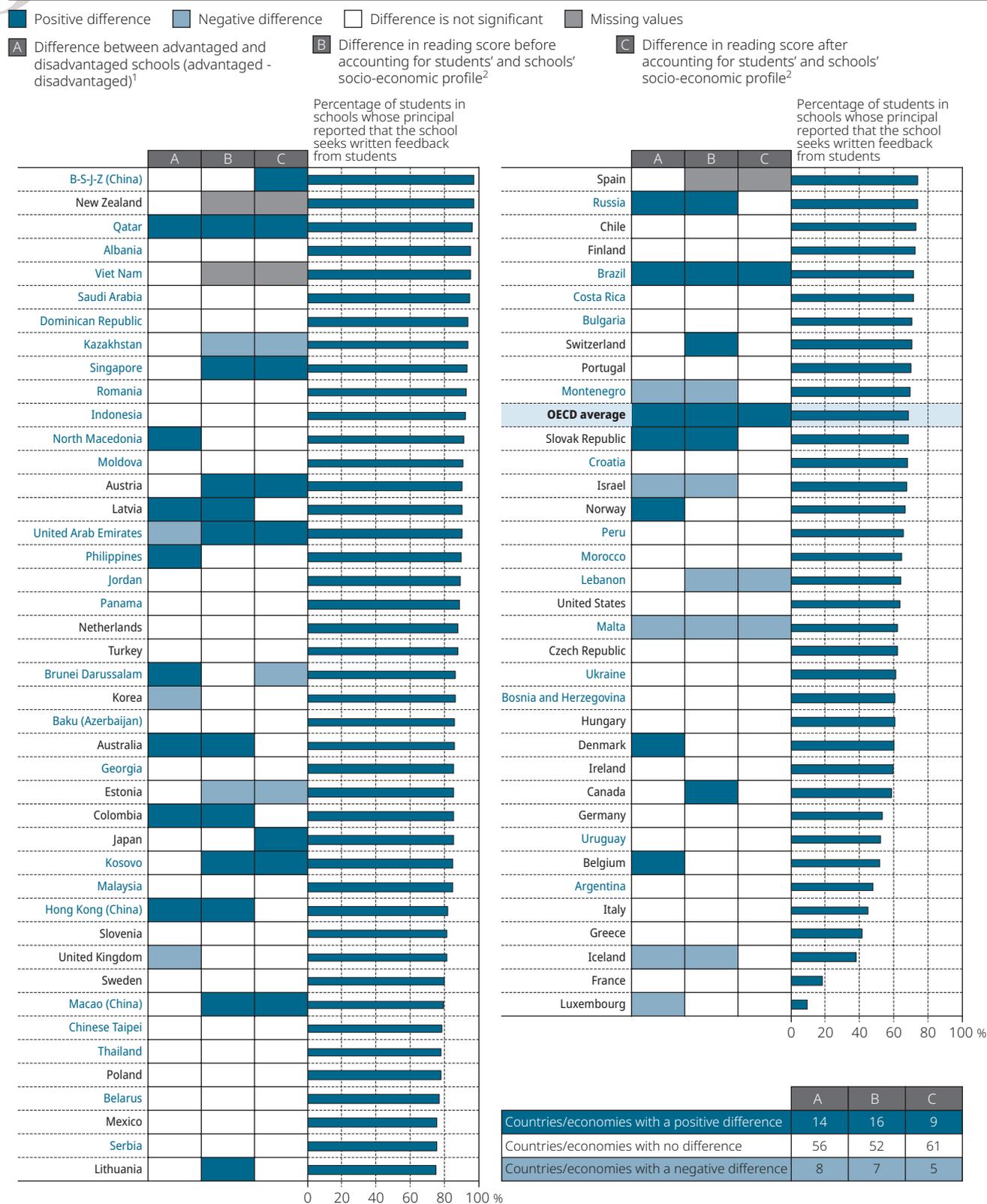
Students in schools that seek feedback from students performed better in reading than students in schools that do not, on average across OECD countries and in 16 countries. Yet, the association between student feedback and student performance is confounded by the fact that, in many education systems, socio-economically advantaged schools tend to seek feedback from their students more than disadvantaged schools do. After accounting for students' and schools' socio-economic profile, seeking feedback from students is still associated with better reading performance, on average across OECD countries and in nine countries and economies. The largest performance differences (at least 25 score points) were observed in B-S-J-Z (China), Macao (China), Qatar, Singapore and the United Arab Emirates.

### Teacher mentoring

On average across OECD countries, 77% of students were enrolled in schools whose principal reported that teacher mentoring is provided in their school. In 62 countries and economies, more than three out of four students were enrolled in schools where teacher mentoring is provided. Teacher mentoring was the least prevalent in Germany, Iceland, Italy and Spain.

On average across OECD countries, teacher mentoring was typically based on the school's own initiative (61% of students were enrolled in such schools), rather than being mandatory (17%). In B-S-J-Z (China), the Czech Republic, Estonia, the Netherlands, Poland and the United Kingdom, at least 85% of students were enrolled in schools whose principal reported that teacher mentoring in their school is provided on the school's initiative. The countries where the availability of mandatory teacher mentoring was the greatest were North Macedonia (84% of students attended such a school), Serbia (70%), Saudi Arabia (69%) and Qatar (66%) (Table V.B1.8.11).

Figure V.8.7 Seeking feedback from students, schools' socio-economic profile and reading performance



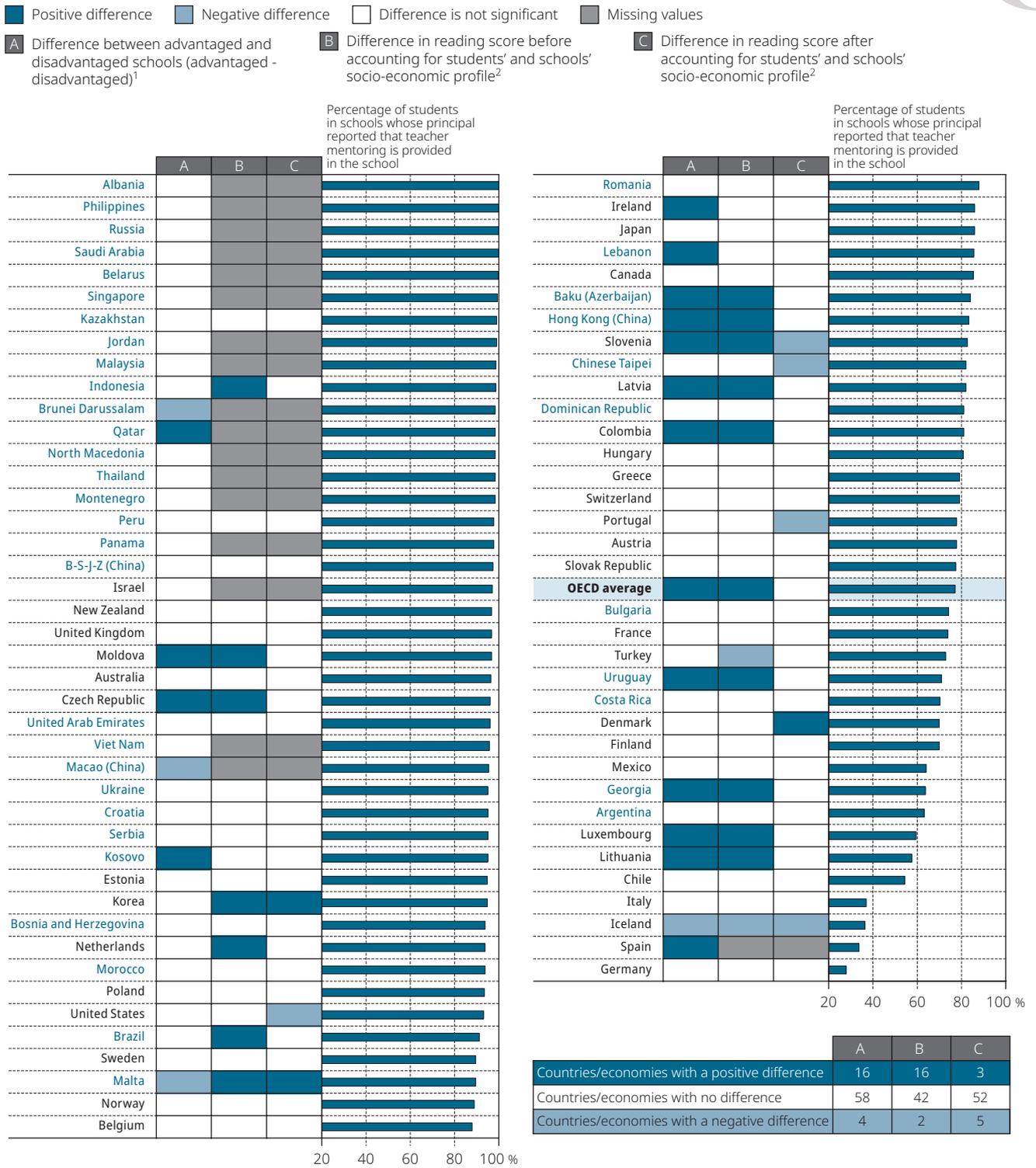
1. A socio-economically disadvantaged (advantaged) school is a school whose socio-economic profile (i.e. the average socio-economic status of the students in the school) is in the bottom (top) quarter of the PISA index of economic, social and cultural status amongst all schools in the relevant country/economy.  
 2. This analysis is restricted to schools with the modal ISCED level for 15-year-old students.

Countries and economies are ranked in descending order of the percentage of students in schools whose principal reported that the school seeks written feedback from students.

Sources: OECD, PISA 2018 Database, Tables V.B1.8.12, V.B1.8.13 and V.B1.8.15.

StatLink <https://doi.org/10.1787/888934131994>

Figure V.8.8 Teacher mentoring, schools' socio-economic profile and reading performance



1. A socio-economically disadvantaged (advantaged) school is a school whose socio-economic profile (i.e. the average socio-economic status of the students in the school) is in the bottom (top) quarter of the PISA index of economic, social and cultural status amongst all schools in the relevant country/economy.  
 2. This analysis is restricted to schools with the modal ISCED level for 15-year-old students.

Countries and economies are ranked in descending order of the percentage of students in schools whose principal reported that teacher mentoring is provided in the school.

Sources: OECD, PISA 2018 Database, Tables V.B1.8.12, V.B1.8.13 and V.B1.8.15.

StatLink <https://doi.org/10.1787/888934132013>

Students in schools where teacher mentoring is provided performed better in reading than students enrolled in schools where no mentoring is provided, on average across OECD countries (a 7 score-point difference), and in 16 countries and economies. Yet, the association between teacher mentoring and student performance is confounded by the fact that, in many education systems, the availability of teacher mentoring is greater in socio-economically advantaged than in disadvantaged schools.

### **HOW THE MEASURES OF EVALUATION AND ASSESSMENT ARE RELATED TO DIFFERENCES IN PERFORMANCE AND EQUITY IN EDUCATION ACROSS COUNTRIES/ECONOMIES (SYSTEM-LEVEL ANALYSIS)**

This section examines how various policies and practices on evaluation and assessment are related to education outcomes at the system level. As shown in Figure V.8.9, two education outcomes are considered: mean performance in reading and equity in reading performance.

At the system level, countries and economies tended to show greater equity in education when they use student assessments to inform parents about their child's progress (Figure V.8.10). Across OECD countries, after accounting for per capita GDP, the higher the percentage of students in schools that use student assessments to inform parents about their child's progress, the weaker the relationship between students' socio-economic status and their reading performance. After accounting for per capita GDP, these correlations are statistically significant in reading, mathematics and science both across OECD countries, and across all countries and economies.<sup>1</sup>

At the system level, countries and economies tended to have better equity in education when they use student assessments to identify aspects of instruction or the curriculum that could be improved (Figure V.8.11). Across OECD countries, after accounting for per capita GDP, the percentage of students in schools that use student assessments to identify aspects of instruction or the curriculum that could be improved was correlated with better equity in performance in reading (partial  $r = 0.41$ ), mathematics (partial  $r = 0.40$ ) and science (partial  $r = 0.45$ ) (Table V.B1.8.16). Across all countries and economies, these correlations were statistically significant after accounting for per capita GDP.<sup>2</sup>

There was no clear pattern in the relationship between reading performance and the prevalence of using student assessments to identify aspects of instruction and the curriculum that could be improved. However, the disparity, related to socio-economic status, in the use of student assessments for this purpose was associated with reading performance. In high-performing education systems, the prevalence of using student assessments for this purpose was similar in socio-economically advantaged and disadvantaged schools or, in some cases, it was more prevalent amongst disadvantaged schools than advantaged schools (Figure V.8.12). For example, in Slovenia, 73% of students in disadvantaged schools attended a school where student assessments are used to identify aspects of instruction and the curriculum that could be improved, while 48% of students in advantaged schools attended such schools. In contrast, in Baku (Azerbaijan), 77% of students in disadvantaged schools attended a school where student assessments are used to identify aspects of instruction and the curriculum that could be improved, while 98% of students in advantaged schools attended such schools (Table V.B1.8.2).

To consider the issue of quality assurance and improvement actions at school, school principals were asked to report whether there are written specifications for student performance at the school and, if so, whether they are mandatory (e.g. based on district or ministry policies) or on the school's initiative. Figure V.8.13 shows that, at the system level, countries and economies tended to show greater equity in education when more students were in schools that have written specifications for student performance based on the school's initiative.<sup>3</sup>

The origin of such written specifications for student performance had a distinct relationship with education outcomes. When written specifications for student performance are mandatory, no clear relationship with equity was observed. However, across all countries and economies, there was a negative relationship with performance in reading, mathematics and science (Table V.B1.8.16). The relationships were significantly negative even after accounting for per capita GDP. This means that lower-performing education systems tended to have more students in schools that have mandatory written specifications for student performance. This could be interpreted to mean that in order to mitigate low performance, district or ministry policies are implemented to set mandatory specifications for student performance.

Another issue regarding quality assurance and improvement actions at school related to equity in performance is whether schools seek written feedback from students (e.g. regarding lessons, teachers or resources). Figure V.8.14 shows that, at the system level, equity in student performance tended to be greater in countries and economies with a higher percentage of students in schools whose principal reported that their school seeks feedback from students (regardless whether such feedback is mandatory or on the school's initiative).<sup>4</sup> Over 12% of the variation in equity in reading performance across OECD countries could be accounted for by differences in the prevalence of schools seeking feedback from students. The correlation between the percentage of students in schools seeking feedback from students and equity in reading, mathematics and science performance was statistically significant, even after accounting for per capita GDP, across OECD countries, and across all countries and economies (Table V.B1.8.16).

Figure V.8.9 [1/4] **Evaluation and assessment, and student performance and equity**

Correlation coefficients between two relevant measures

		OECD countries			
		Mean reading score		Equity in reading	
		Before accounting for per capita GDP	After accounting for per capita GDP	Before accounting for per capita GDP	After accounting for per capita GDP
Purposes of student assessment	To guide students' learning			<i>0.32</i>	<i>0.32</i>
	To guide students' learning: top-bottom quarter difference in schools' socio-economic profile				
	To inform parents about their child's progress			<i>0.30</i>	<i>0.31</i>
	To inform parents about their child's progress: top-bottom quarter difference in schools' socio-economic profile				
	To make decisions about students' retention or promotion				<i>-0.29</i>
	To make decisions about students' retention or promotion: top-bottom quarter difference in schools' socio-economic profile				
	To group students for instructional purposes				
	To group students for instructional purposes: top-bottom quarter difference in schools' socio-economic profile				
	To compare school performance to district/national performance				
	To compare school performance to district/national performance: top-bottom quarter difference in schools' socio-economic profile				
	To monitor the school's progress from year to year				
	To monitor the school's progress from year to year: top-bottom quarter difference in schools' socio-economic profile				
	To make judgements about teachers' effectiveness				
	To make judgements about teachers' effectiveness: top-bottom quarter difference in schools' socio-economic profile	<i>-0.44</i>	<i>-0.39</i>		
	To identify aspects of instruction/curriculum be improved			<i>0.39</i>	<i>0.41</i>
	To identify aspects of instruction/curriculum be improved: top-bottom quarter difference in schools' socio-economic profile	<i>-0.32</i>	<i>-0.30</i>		
	To adapt teaching to the students' needs			<i>0.41</i>	<i>0.41</i>
	To adapt teaching to the students' needs: top-bottom quarter difference in schools' socio-economic profile			<i>0.35</i>	<i>0.35</i>
	To adapt teaching to compare the school with other schools				
	To adapt teaching to compare the school with other schools: top-bottom quarter difference in schools' socio-economic profile				
To award certificates to students					
To award certificates to students: top-bottom quarter difference in schools' socio-economic profile					
School accountability	Achievement data are posted publicly				
	Achievement data are tracked over time by an administrative authority				
	Achievement data are provided directly to parents	<i>-0.34</i>			

**Notes:** Correlation coefficients range from -1.00 (i.e. a perfect negative linear association) to +1.00 (i.e. a perfect positive linear association). When a correlation coefficient is 0, there is no linear relationship between the two measures.

Only statistically significant coefficients are shown. Values that are statistically significant at the 10% level ( $p < 0.10$ ) are in italics. All other values are statistically significant at the 5% level ( $p < 0.05$ ).

**Source:** OECD, PISA 2018 Database, Table V.B1.8.16.

**StatLink**  <https://doi.org/10.1787/888934132032>

Figure V.8.9 [2/4] **Evaluation and assessment, and student performance and equity**

Correlation coefficients between two relevant measures

		OECD countries			
		Mean reading score		Equity in reading	
		Before accounting for per capita GDP	After accounting for per capita GDP	Before accounting for per capita GDP	After accounting for per capita GDP
Quality assurance and improvement actions at school	Internal evaluation/self-evaluation				
	Internal evaluation/self-evaluation is mandatory			<i>0.30</i>	<i>0.30</i>
	Internal evaluation/self-evaluation is on the school's initiative				
	There is no internal evaluation/self-evaluation				
	External evaluation				
	External evaluation is mandatory				
	External evaluation is on the school's initiative				
	There is no external evaluation				
	Written specification of school's curricular profile and education goals				
	Written specification of school's curricular profile and education goals is mandatory			<i>-0.30</i>	<i>-0.30</i>
	Written specification of school's curricular profile and education goals is on the school's initiative			<i>0.35</i>	<i>0.35</i>
	There is no written specification of school's curricular profile and education goals				
	Written specification of student performance standards				
	Written specification of student performance standards is mandatory				
	Written specification of student performance standards is on the school's initiative			<i>0.42</i>	<i>0.42</i>
	There is no written specification of student performance standards				
	Systematic recording of data				
	Systematic recording of data is mandatory				
	Systematic recording of data is on the school's initiative				
	There is no systematic recording of data				
	Systematic recording of student test results and graduation rates (combined)				
	Systematic recording of student test results and graduation rates is mandatory				
	Systematic recording of student test results and graduation rates is on the school's initiative				
	There is no systematic recording of student test results and graduation rates				
	Seeking written feedback from students		<i>0.33</i>	<i>0.36</i>	<i>0.39</i>
	Seeking written feedback from students is mandatory			<i>0.31</i>	<i>0.31</i>
	Seeking written feedback from students is on the school's initiative		<i>0.31</i>		
	There is no written feedback from students		<i>-0.33</i>	<i>-0.36</i>	<i>-0.39</i>
	Teacher mentoring	<i>0.35</i>	<i>0.40</i>		
	Teacher mentoring is mandatory				
	Teacher mentoring is on the school's initiative	<i>0.40</i>	<i>0.45</i>		
	There is no teacher mentoring	<i>-0.35</i>	<i>-0.40</i>		
	Regular consultation on school improvement at least over 6 months				
Regular consultation on school improvement at least over 6 months is mandatory			<i>0.30</i>	<i>0.31</i>	
Regular consultation on school improvement at least over 6 months is on the school's initiative					
There is no regular consultation on school improvement at least over 6 months					
Implementation of a standardised policy for reading subjects					
Implementation of a standardised policy for reading subjects is mandatory					
Implementation of a standardised policy for reading subjects is on the school's initiative					
There is no implementation of a standardised policy for reading subjects					

**Notes:** Correlation coefficients range from -1.00 (i.e. a perfect negative linear association) to +1.00 (i.e. a perfect positive linear association). When a correlation coefficient is 0, there is no linear relationship between the two measures.

Only statistically significant coefficients are shown. Values that are statistically significant at the 10% level ( $p < 0.10$ ) are in italics. All other values are statistically significant at the 5% level ( $p < 0.05$ ).

**Source:** OECD, PISA 2018 Database, Table V.B1.8.16.

**StatLink**  <https://doi.org/10.1787/888934132032>

Figure V.8.9 [3/4] **Evaluation and assessment, and student performance and equity**

Correlation coefficients between two relevant measures

		All countries and economies			
		Mean reading score		Equity in reading	
		Before accounting for per capita GDP	After accounting for per capita GDP	Before accounting for per capita GDP	After accounting for per capita GDP
Purposes of student assessment	To guide students' learning		-0.23		
	To guide students' learning: top-bottom quarter difference in schools' socio-economic profile	-0.21	-0.20		
	To inform parents about their child's progress		-0.24	0.30	0.29
	To inform parents about their child's progress: top-bottom quarter difference in schools' socio-economic profile	-0.20			
	To make decisions about students' retention or promotion	-0.28	-0.21		
	To make decisions about students' retention or promotion: top-bottom quarter difference in schools' socio-economic profile	-0.24			
	To group students for instructional purposes		-0.37	0.22	
	To group students for instructional purposes: top-bottom quarter difference in schools' socio-economic profile				
	To compare school performance to district/national performance				
	To compare school performance to district/national performance: top-bottom quarter difference in schools' socio-economic profile				
	To monitor the school's progress from year to year	-0.31	-0.29		0.20
	To monitor the school's progress from year to year: top-bottom quarter difference in schools' socio-economic profile				
	To make judgements about teachers' effectiveness	-0.40	-0.40	0.20	0.22
	To make judgements about teachers' effectiveness: top-bottom quarter difference in schools' socio-economic profile	-0.35			
	To identify aspects of instruction/curriculum be improved	-0.20	-0.23		0.20
	To identify aspects of instruction/curriculum be improved: top-bottom quarter difference in schools' socio-economic profile	-0.38	-0.29		
	To adapt teaching to the students' needs				
	To adapt teaching to the students' needs: top-bottom quarter difference in schools' socio-economic profile	-0.24			
	To adapt teaching to compare the school with other schools	-0.21	-0.21	0.20	0.20
	To adapt teaching to compare the school with other schools: top-bottom quarter difference in schools' socio-economic profile				
To award certificates to students		-0.22	0.27	0.26	
To award certificates to students: top-bottom quarter difference in schools' socio-economic profile					
School accountability	Achievement data are posted publicly			0.19	
	Achievement data are tracked over time by an administrative authority	-0.34	-0.31		
	Achievement data are provided directly to parents	-0.44	-0.30		

**Notes:** Correlation coefficients range from -1.00 (i.e. a perfect negative linear association) to +1.00 (i.e. a perfect positive linear association). When a correlation coefficient is 0, there is no linear relationship between the two measures.

Only statistically significant coefficients are shown. Values that are statistically significant at the 10% level ( $p < 0.10$ ) are in italics. All other values are statistically significant at the 5% level ( $p < 0.05$ ).

**Source:** OECD, PISA 2018 Database, Table V.B1.8.16.

**StatLink**  <https://doi.org/10.1787/888934132032>

Figure V.8.9 [4/4] **Evaluation and assessment, and student performance and equity**

Correlation coefficients between two relevant measures

		All countries and economies			
		Mean reading score		Equity in reading	
		Before accounting for per capita GDP	After accounting for per capita GDP	Before accounting for per capita GDP	After accounting for per capita GDP
Quality assurance and improvement actions at school	Internal evaluation/self-evaluation				
	Internal evaluation/self-evaluation is mandatory			<i>0.19</i>	<i>0.20</i>
	Internal evaluation/self-evaluation is on the school's initiative				
	There is no internal evaluation/self-evaluation				
	External evaluation				
	External evaluation is mandatory				
	External evaluation is on the school's initiative	-0.24	-0.20		
	There is no external evaluation				
	Written specification of school's curricular profile and education goals				
	Written specification of school's curricular profile and education goals is mandatory	-0.24			
	Written specification of school's curricular profile and education goals is on the school's initiative	<i>0.21</i>			
	There is no written specification of school's curricular profile and education goals				
	Written specification of student performance standards	-0.31	-0.31		
	Written specification of student performance standards is mandatory	-0.35	-0.23		
	Written specification of student performance standards is on the school's initiative			0.34	0.34
	There is no written specification of student performance standards	0.31	0.31		
	Systematic recording of data	-0.25	-0.26		
	Systematic recording of data is mandatory	-0.22	-0.22		
	Systematic recording of data is on the school's initiative				
	There is no systematic recording of data	0.25	0.26		
	Systematic recording of student test results and graduation rates (combined)			<i>0.21</i>	<i>0.22</i>
	Systematic recording of student test results and graduation rates is mandatory	-0.29	-0.30		
	Systematic recording of student test results and graduation rates is on the school's initiative	0.24	0.25		
	There is no systematic recording of student test results and graduation rates			-0.21	-0.22
	Seeking written feedback from students			<i>0.22</i>	<i>0.23</i>
	Seeking written feedback from students is mandatory	-0.24	-0.31		
	Seeking written feedback from students is on the school's initiative				
	There is no written feedback from students			-0.22	-0.23
	Teacher mentoring				
	Teacher mentoring is mandatory	-0.51	-0.50		
	Teacher mentoring is on the school's initiative	0.40	0.37		
	There is no teacher mentoring				
Regular consultation on school improvement at least over 6 months	-0.34	-0.45	<i>0.20</i>	<i>0.20</i>	
Regular consultation on school improvement at least over 6 months is mandatory	-0.48	-0.57			
Regular consultation on school improvement at least over 6 months is on the school's initiative			<i>0.20</i>	<i>0.20</i>	
There is no regular consultation on school improvement at least over 6 months	0.34	0.45	-0.20	-0.20	
Implementation of a standardised policy for reading subjects	-0.41	-0.44			
Implementation of a standardised policy for reading subjects is mandatory	-0.50	-0.49			
Implementation of a standardised policy for reading subjects is on the school's initiative					
There is no implementation of a standardised policy for reading subjects	0.41	0.44			

**Notes:** Correlation coefficients range from -1.00 (i.e. a perfect negative linear association) to +1.00 (i.e. a perfect positive linear association). When a correlation coefficient is 0, there is no linear relationship between the two measures.

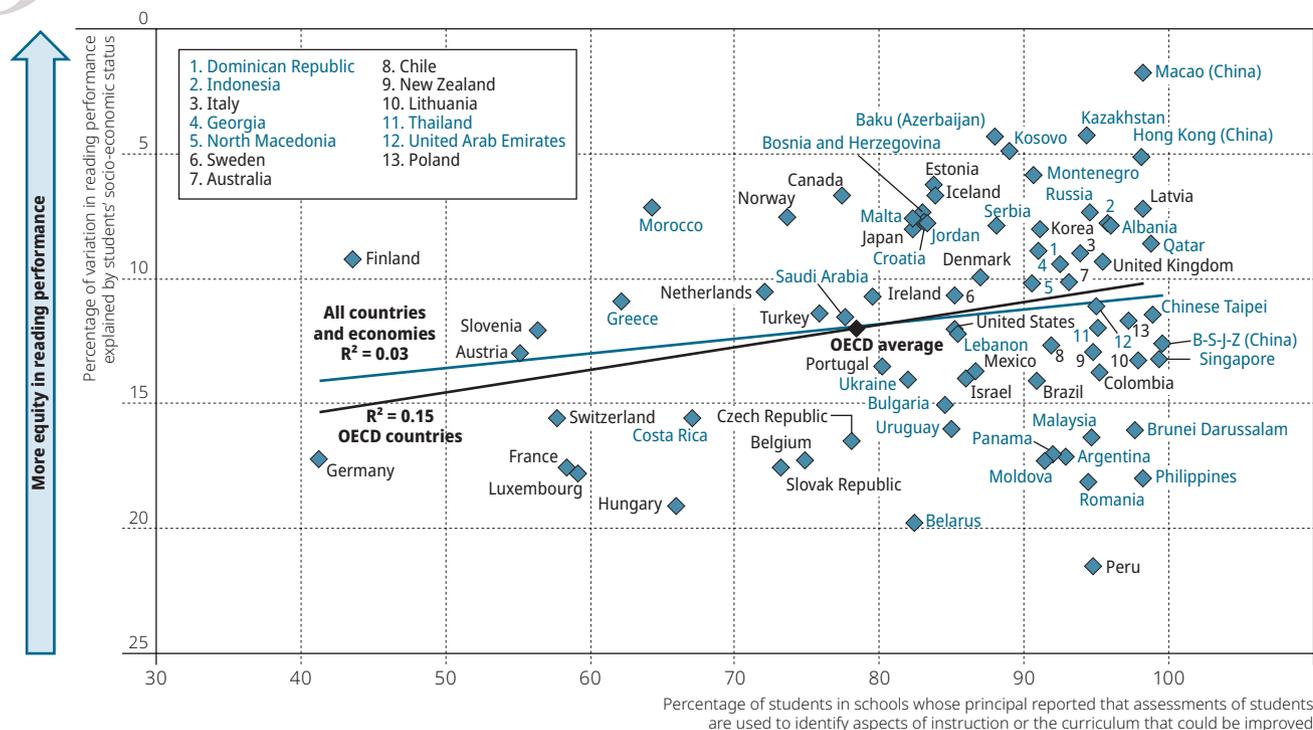
Only statistically significant coefficients are shown. Values that are statistically significant at the 10% level ( $p < 0.10$ ) are in italics. All other values are statistically significant at the 5% level ( $p < 0.05$ ).

**Source:** OECD, PISA 2018 Database, Table V.B1.8.16.

**StatLink**  <https://doi.org/10.1787/888934132032>



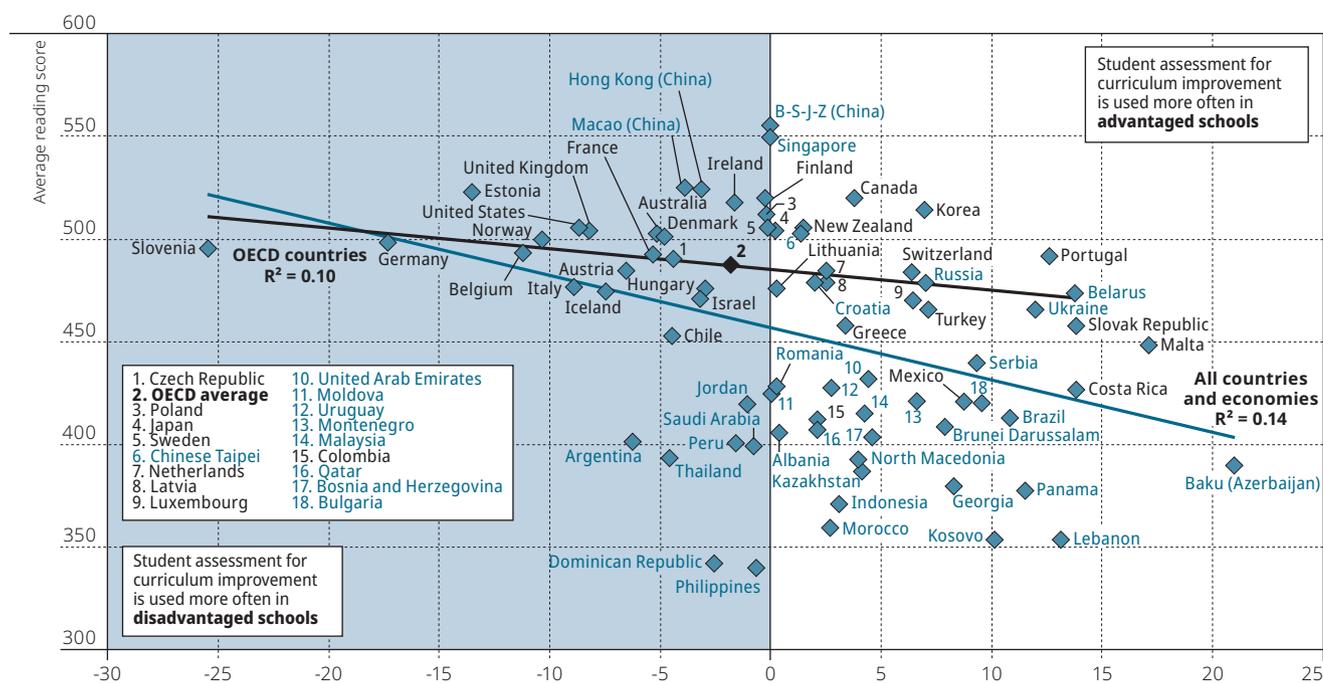
Figure V.8.11 Using student assessment for curriculum improvement and equity in reading performance



Sources: OECD, PISA 2018 Database, Tables II.B1.2.3 and V.B1.8.1.

StatLink <https://doi.org/10.1787/888934132070>

Figure V.8.12 Socio-economic disparities in using student assessment for curriculum improvement and mean reading performance



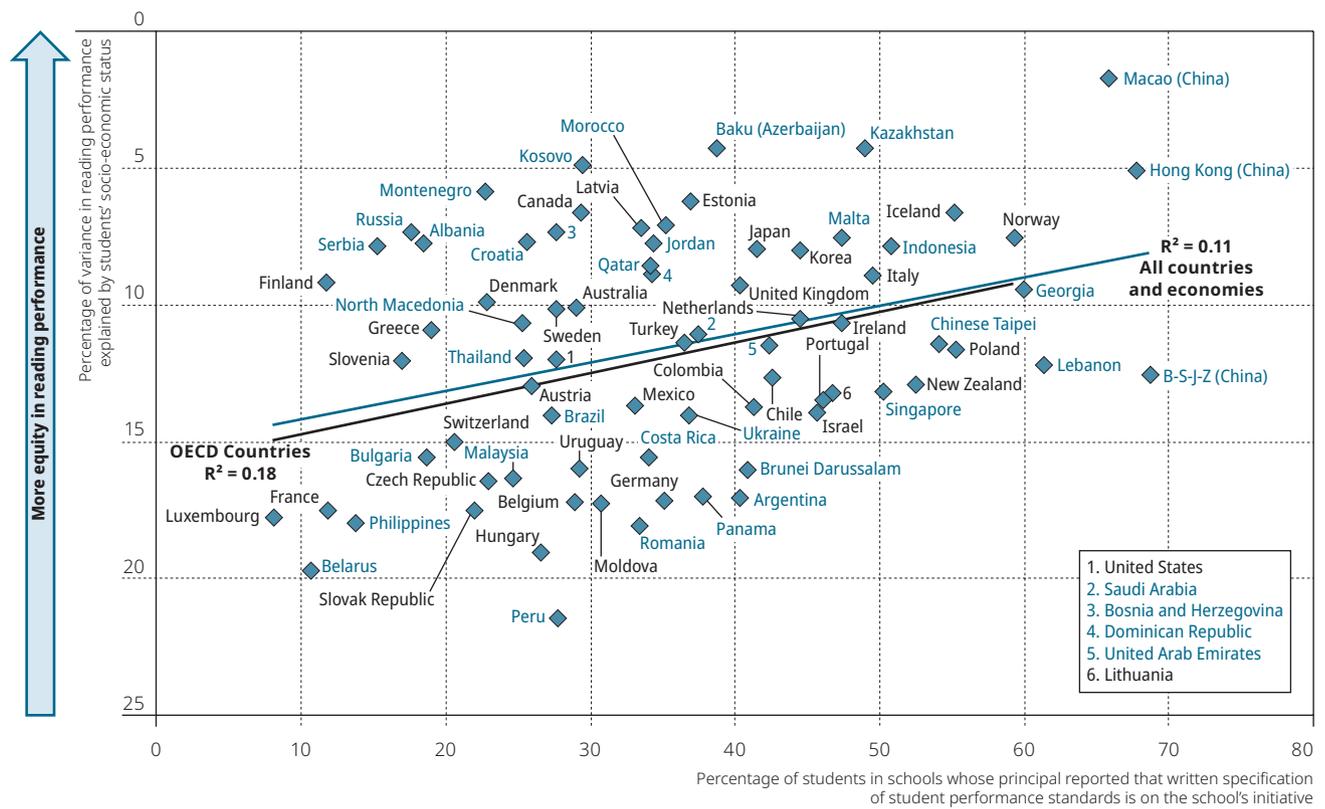
Socio-economic disparities in using student assessment for curriculum improvement: Difference between advantaged and disadvantaged schools

Note: A socio-economically disadvantaged (advantaged) school is a school whose socio-economic profile (i.e. the average socio-economic status of the students in the school) is in the bottom (top) quarter of the PISA index of economic, social and cultural status amongst all schools in the relevant country/economy.

Sources: OECD, PISA 2018 Database, Tables I.B1.4 and V.B1.8.2.

StatLink <https://doi.org/10.1787/888934132089>

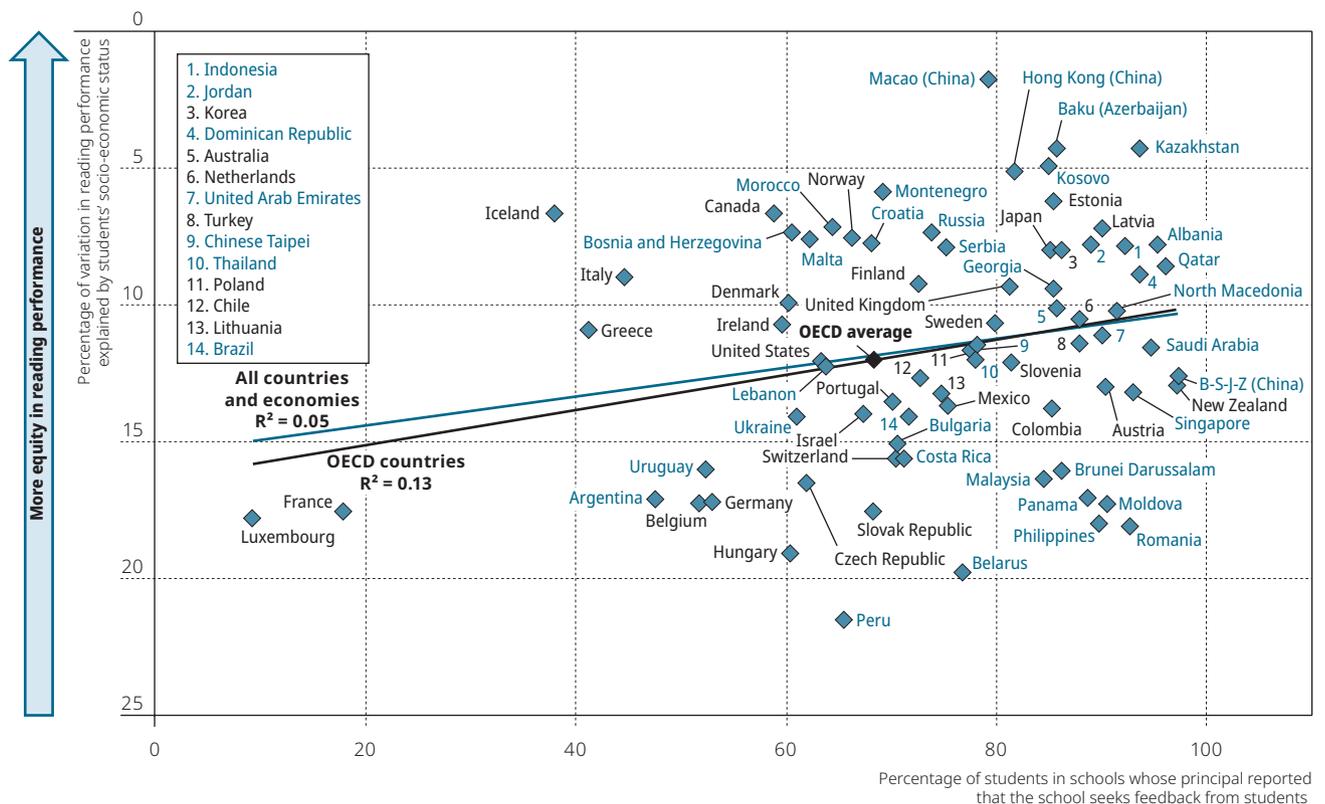
Figure V.8.13 **Written specification of student performance standards on the school's initiative and equity in reading performance**



Sources: OECD, PISA 2018 Database, Tables II.B1.2.3 and V.B1.8.11.

StatLink <https://doi.org/10.1787/888934132108>

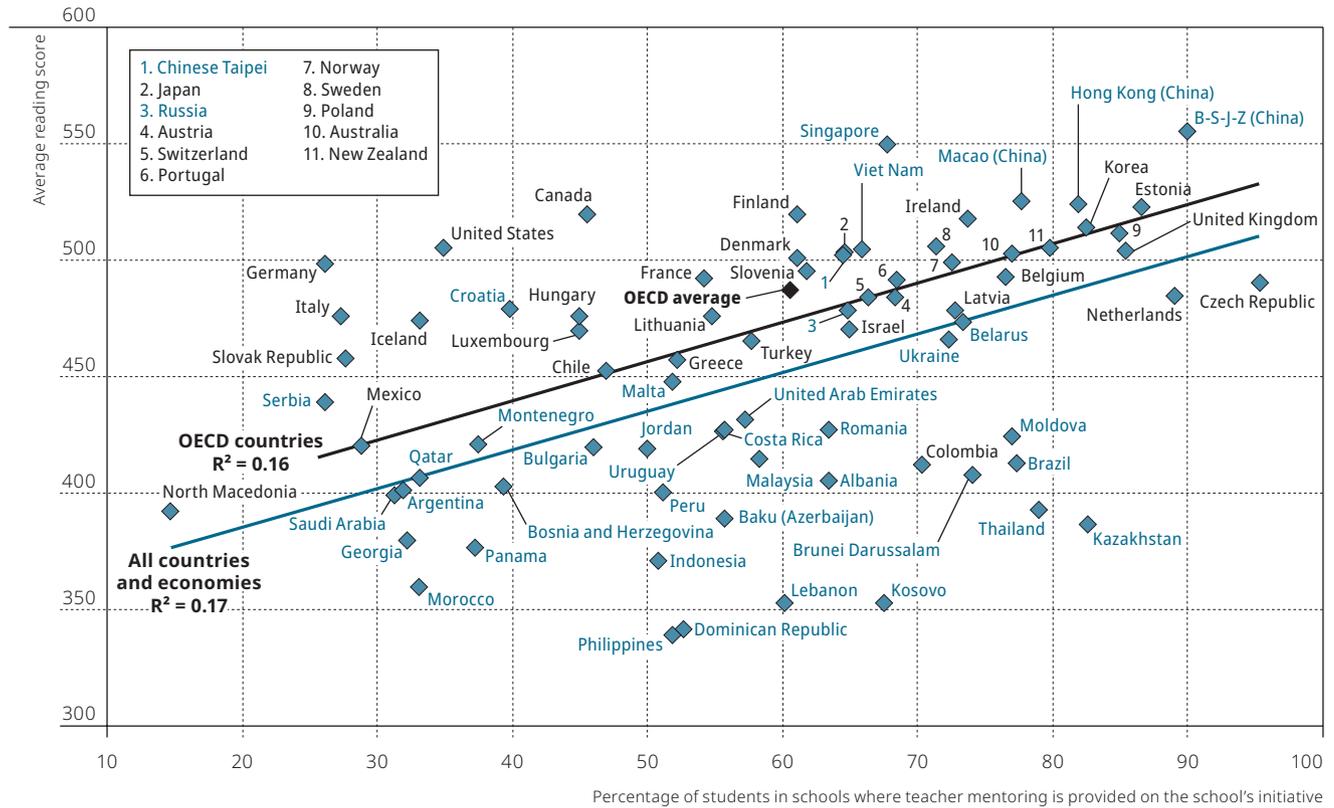
Figure V.8.14 **Seeking feedback from students, and equity in reading performance**



StatLink <https://doi.org/10.1787/888934132127>

Figure V.8.15 Teacher mentoring on the school's initiative and reading performance

Based on principals' reports



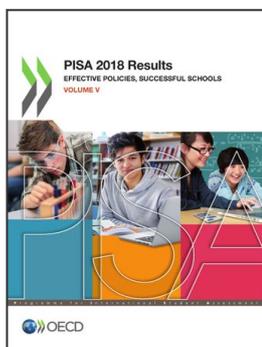
StatLink <https://doi.org/10.1787/888934132146>

## Notes

1. After excluding low-performing countries/economies (i.e. mean performance in reading lower than 413 points), the strength of the association across OECD countries remained almost unaltered (after exclusion,  $R^2 = 0.29$ ), whereas across all countries/economies, the association strengthened (after exclusion,  $R^2 = 0.37$ ).
2. After excluding low-performing countries/economies (i.e. mean performance in reading lower than 413 points), the strength of the association increased both across OECD countries (after exclusion,  $R^2 = 0.42$ ), and across all countries/economies (after exclusion,  $R^2 = 0.30$ ).
3. After excluding low-performing countries/economies (i.e. mean performance in reading lower than 413 points), the strength of the association increased both across OECD countries (after exclusion,  $R^2 = 0.43$ ), and across all countries/economies (after exclusion,  $R^2 = 0.40$ ).
4. After excluding low-performing countries/economies (i.e. mean performance in reading lower than 413 points), the strength of the association across OECD countries remained almost unaltered (after exclusion,  $R^2 = 0.37$ ), whereas across all countries/economies, the association was not significant.

## References

- Förster, C.** et al. (2017), *El poder de la evaluación en el aula*. [8]
- Hattie, J.** and **H. Timperley** (2007), "The Power of Feedback", *Review of Educational Research*, doi: 10.3102/003465430298487, pp. 81-112, <http://dx.doi.org/10.3102/003465430298487>. [6]
- Lesaux, N.** (2006), "Meeting Expectations? An Empirical Investigation of a Standards-Based Assessment of Reading Comprehension", *Educational Evaluation and Policy Analysis - EDUC EVAL POLICY ANAL*, Vol. 28, pp. 315-333, <http://dx.doi.org/10.3102/01623737028004315>. [3]
- Looney, J.** (2011), "Alignment in Complex Education Systems: Achieving Balance and Coherence", *OECD Education Working Papers*, No. 64, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5kg3vg5lx8r8-en>. [4]
- OECD** (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264190658-en>. [1]
- OECD** (2013), "The evaluation and assessment framework: Embracing a holistic approach", in *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264190658-6-en>. [7]
- Rosenkvist, M.** (2010), "Using Student Test Results for Accountability and Improvement: A Literature Review", *OECD Education Working Papers*, No. 54, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5km4htwzvbv30-en>. [2]
- Shepard, L.** (2000), "The Role of Assessment in a Learning Culture", *Educational Researcher*, doi: 10.3102/0013189X029007004, pp. 4-14, <http://dx.doi.org/10.3102/0013189X029007004>. [5]



**From:**  
**PISA 2018 Results (Volume V)**  
Effective Policies, Successful Schools

**Access the complete publication at:**  
<https://doi.org/10.1787/ca768d40-en>

**Please cite this chapter as:**

OECD (2020), "Assessments, evaluations and how they are used", in *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/cd754fae-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

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

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