TALIS

Positive, High-achieving Students?

What Schools and Teachers Can Do

Foreword

Large-scale education surveys are a core asset for education policy making. They foster new analyses and findings about ways to improve the professionalism of teachers and school leaders, and the quality and equity of education around the world. The combination of several trends, including the increasing migration of surveys to digital devices and the technical progress made in data collection, storage and processing, are leading to unprecedented opportunities in the education sphere. Over the past two decades, the OECD Directorate for Education and Skills, together with its member and partner countries, has collected huge volumes of valid, reliable, comparable, rigorous, interpretable and policy-relevant data on teachers, school leaders and students around the world. Let us take a moment to zoom in on the two largest OECD education surveys to date.

The OECD Programme for International Student Assessment (PISA) provides the most comprehensive and rigorous international assessment of student learning outcomes to date. PISA also collects large amounts of data on students' family backgrounds, schools, learning conditions and on students' lives in general. Results from PISA indicate the quality and equity of cognitive and social-emotional outcomes attained around the world, and allow educators and policy makers to learn from the policies and practices applied in other countries. Today, PISA brings together more than 90 countries and economies, representing 80% of the world economy, in a global conversation about education.

The OECD Teaching and Learning International Survey (TALIS) is the largest international and periodic survey asking teachers and school leaders about their working conditions and learning environments. Based on the voices of more than 260 000 teachers representing more than 8 million teachers from around 50 countries and economies worldwide, TALIS 2018 offers orientations to help strengthen the knowledge and skills of the teaching workforce to support teaching and learning.

Now, imagine combining all these data into one, big dataset. Here is what happened: in 2018, nine countries and economies administered both TALIS and PISA to teachers, school leaders and students of the same schools. This has resulted in the creation of the TALIS-PISA link dataset, which comprises thousands of variables from more than 30 000 students and more than 15 000 teachers from more than 1 000 schools on four different continents. The OECD's job is to make sense of it all: to understand "what the data say" and extract important patterns that can inform education policy.

To achieve this, analysis relies on the thorough conceptual process that has taken place prior to collecting this data. None of the variables in the TALIS and PISA datasets were included by chance! Every question in our questionnaires and assessments was devised, revised and refined by large international expert groups following a comprehensive iterative process; informed by previous research findings, conceptual frameworks, outcomes of survey piloting and field trialling; and in consultation with many stakeholders from the teaching profession and the education sector. All these data were collected with the view of shedding light on important education issues.

At the same time, the challenges related to the complexity and size of these TALIS-PISA link data required complementing this theoretical approach with data-driven techniques borrowed from the rapidly expanding field of machine or statistical learning. At the OECD, not only do we examine the role that artificial intelligence (AI) might have in improving education processes and preparing students for increasingly automated economies and societies – readers can look at our paper on *Trustworthy artificial intelligence (AI) in education* in support of the G20 artificial intelligence dialogue, for example – but we are also agents of this transformation. This report is the first of its kind to explore the potential of one particular form of AI – supervised statistical learning – to improve our understanding of the nexus between teaching and learning.

Yet, while applying statistical learning to the TALIS-PISA link data, the analysis never loses sight of theory and inputs from education stakeholders. This report maintains a steady dialogue between theory, fieldwork and data. It first addresses three questions: According to past research, what teacher and school factors matter the most for student learning outcomes? For student social-emotional development? For all students to thrive and fulfil their potential, regardless of their personal characteristics? Then the report turns to the data and tries to understand "what the data say" using statistical learning complemented with more standard statistical analyses. The analysis goes back and forth between data results and the available research literature. New statistical analyses can lead to unexpected findings. Yet, existing research studies can be used to validate, complement and put the new findings into perspective.

So what have we learned? If there was only one conclusion to take away from this report, it is that what teachers do in and outside the classroom matters the most – and the most directly – for the cognitive and social-emotional outcomes of the school's students. Classroom practices that create opportunities to learn, teachers' use of working time, as well as the well-being and job satisfaction of the teachers are among the most influential school factors. But this report does not only confirm the crucial role of teachers in young people's development. It also sheds light on other actors. Students' classmates and schoolmates, as well as the school's culture and leadership (including the role parents play) are also found to matter a lot for student outcomes. All this suggests that any education reform or any change in education practice should tackle one or several of these dimensions first for students to be able to thrive and be active participants in our economies and societies.

A few months ago, we released the OECD *Global Teaching Insights*, our first large-scale video study of teaching, which provides a detailed account of how teaching and learning take place in the classrooms. For the first time, the OECD has implemented a study that combines the observation of lessons via video and a longitudinal design able to measure growth in student outcomes. With this report, *Positive, High-achieving Students? What Schools and Teachers Can Do*, the OECD is now trialling new research methods to shed light on the key factors that can be activated to raise student outcomes and reduce disparities deriving from personal characteristics and circumstances. And the OECD will continue to do so in the future. We will keep reflecting on, revising and refining our research and policy questions, our study designs and our research methods to reap the most of what teachers, school leaders and students tell us about teaching and learning. Stay tuned.

Andreas Schleicher,

Director for Education and Skills

Special Advisor on Education Policy to the OECD Secretary-General

Acknowledgements

This report is the outcome of a collaboration among the countries and economies participating in TALIS and those participating in PISA, the OECD Secretariat, the European Commission and the TALIS and PISA international consortia.

The development of this report was guided by Andreas Schleicher, Yuri Belfali and Karine Tremblay (TALIS Project Manager). Noémie Le Donné led its development and production, with contributions from Gabor Fülöp and Emily Groves. Noémie Le Donné was the lead author of Chapters 1 and 3. Gabor Fülöp was the lead author of Chapter 2. He also authored Chapter 4, with contributions from Massimo Loi and Noémie Le Donné, as well as the annexes.

All statistical analyses were performed by Gabor Fülöp, with contributions from Massimo Loi. Maxence Castiello and Hélène Guillou provided support for statistical editing.

Emily Groves supported report preparation, production, project co-ordination and communications. Cassandra Davis and Henri Pearson provided support for report production and communications. Clara Young was the main editor of the report, with contributions from Emily Groves and Henri Pearson.

The authors wish to thank Francesco Avvisati, Yuri Belfali, Pablo Fraser, Pauline Givord, Aakriti Kalra, Andreas Schleicher, Javier Suarez Alvarez and Karine Tremblay, as well as members of the TALIS Governing Board and the TALIS national project managers, in particular, those from the nine countries/economies participating in the TALIS-PISA link option, who all provided valuable feedback and input at various stages of the data analysis and report production. The development of the report was steered by the TALIS Governing Board, chaired by João Costa (Portugal).

Table of contents

Tables

Figures

Boxes

Reader's guide

The results referred to in this volume are provided in Annex C via links to online tables.

Country coverage

The Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link, referred to as the TALIS-PISA link in the report, covers schools that participated in both TALIS and PISA. TALIS-PISA link 2018 features results on students, teachers and school principals from nine countries and economies: Australia, Ciudad Autónoma de Buenos Aires (referred to as CABA [Argentina]), Colombia, the Czech Republic, Denmark, Georgia, Malta, Turkey, Viet Nam.

Reporting student data

The report uses "students" as shorthand for the TALIS-PISA link target population of 15-year-old students who are surveyed by PISA. In PISA 2018, the international target population of students includes those who are aged between 15 years and 3 months and 16 years and 2 months at the time of the 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. Since the target population of students is defined by age, not by grade level, 15-year-old students surveyed by PISA are enrolled in a programme either at International Standard Classification of Education (ISCED) 2 level (lower secondary education) or at ISCED 3 level (upper secondary education). In the case of student-level analyses, students' outcomes and characteristics are weighted by the student weights. In the case of school-level analyses, students' answers are averaged at the school level and are weighted by the school weights.

Reporting teacher data

The report uses "teachers" as shorthand for the TALIS-PISA link target population of teachers who teach in schools surveyed by PISA and who, as part of their regular duties in a target school, provide instruction for 15-year-old students in programmes at ISCED 2 level (lower secondary education) or at ISCED 3 level (upper secondary education), or at both levels.

As the TALIS-PISA link data only permit the matching of a sample of teachers teaching 15-year-old students in a school with a sample of 15-year-old students of that same school, information on teachers is averaged at the school level. In the case of student-level analyses, teachers' responses are weighted by the student weights, while, in the case of school-level analyses, they are weighted by the school weights.

Reporting principal data

The report uses “principals” and “school leaders” as equivalent shorthand for the TALIS-PISA link target population of principals working in schools surveyed by PISA that provide instruction for 15-year-old students in programmes at ISCED 2 level (lower secondary education) or at ISCED 3 level (upper secondary education) or at both levels. In the case of student-level analyses, principals’ responses are weighted by the student weights, while in the case of school-level analyses, they are weighted by the school weights.

International averages

The TALIS-PISA link average corresponds to the arithmetic mean of the respective country estimates with available data. In Chapter 2, the TALIS-PISA link average covers all participating countries and economies, excluding Viet Nam.[1] In Chapter 3, the TALIS-PISA link average covers all participating countries and economies. In Chapter 4, the TALIS-PISA link average excludes Viet Nam in the case of student-level analyses and it excludes Malta[2] and Viet Nam in the case of school-level analyses.

In the case of some countries, data may not be available for specific indicators, or specific categories may not apply. Therefore, readers should keep in mind that the terms “TALIS-PISA link average” refers to the countries included in the respective averages. Each of these averages may not necessarily be consistent across all columns of a table.

Data underlying the figures

One symbol is used to denote non-reported estimates:

- n: Data is missing due to computational process (e.g. regression coefficient is missing if an independent variable in a regression model is omitted due to collinearity).

TALIS data are based exclusively on self-reports from teachers and school leaders and, therefore, represent their opinions, perceptions, beliefs and accounts of their activities. No data imputation from administrative data or other studies is conducted and, as with any self-reported data, this information is subjective and may, therefore, differ from data collected through other means (e.g. administrative data or video observations). The same is true of school leaders’ reports about school characteristics and practices, which may differ from descriptions provided by administrative data at a national or local government level.

PISA data focuses on the extent to which 15-year-old students have acquired key competencies essential for full participation in social and economic life. PISA assessments do not just ascertain whether students can reproduce what they have learned, but they also examine how well students can extrapolate from what they have learned and apply their knowledge. Hence, while PISA focusses on students’ competences and on how well these competences are applied in different contexts, it may not reflect the curriculum for 15-year-old students. In each round of PISA, one subject among reading, mathematics and science, is tested in detail. The main subject in 2018 was reading.

Caution is warranted when interpreting the results presented in this report. Due to the specific survey design of the TALIS-PISA link data:

- Findings only reflect the overall relationships between a student and a school’s average teacher, since no direct link can be drawn between a teacher and his/her students.
- Results cannot be interpreted as causal, but only as correlational, given the cross-sectional nature of the data and the complexity and reciprocity of the relationships analysed.

- Generalisability of the findings is limited due to the limited number of countries/economies included in the TALIS-PISA link.

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 (SE) in this publication have been rounded to one, two or three decimal places. Where the value 0.0, 0.00 or 0.000 is shown, this does not imply that the standard error is zero, but that it is smaller than 0.05, 0.005 or 0.0005, respectively.

Focusing on statistically significant differences

This volume only comments on statistically significant estimates, with the significance level set to 5%, unless otherwise specified. These are denoted in darker colours in figures and in bold font in tables. See Annex B for further information.

Further technical documentation

For further information on TALIS and PISA documentation, instruments and methodology, see the *TALIS 2018 Technical Report* (OECD, 2019[1]), *TALIS 2018 and TALIS Starting Strong 2018 User Guide* (OECD, 2019[2]) and the *PISA 2018 Technical Report* (OECD, 2020[3]).

This report uses the OECD StatLinks service. All tables and charts are assigned a URL leading to a corresponding Excel™ workbook containing the underlying data. These URLs are stable and will remain unchanged over time. In addition, readers of the e-books will be able to click directly on these links and the workbook will open in a separate window if their Internet browser is open and running.

References

OECD (2020), *PISA 2018 Technical Report*, https://www.oecd.org/pisa/data/pisa2018technicalreport/ (accessed on 16 November 2020). [3]

OECD (2019), *TALIS 2018 and TALIS Starting Strong 2018 User Guide*, OECD, Paris, http://www.oecd.org/education/talis/TALIS_2018-TALIS_Starting_Strong_2018_User_Guide.pdf. [2]

OECD (2019), *TALIS 2018 Technical Report*, OECD, Paris, http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf. [1]

Notes

[1] Since Viet Nam does not have data on PISA test scores, it is not included in the analyses presented in Chapters 2 and 4.

[2] In Malta, there are only 17 out of the 44 schools that are not single-gender schools (i.e. all students surveyed in the school are same-gender students) and where the within-school differences in performance between girls and boys can be computed. Thus, it is not included in the school-level analysis presented in Chapter 4.

Executive summary

There are many things we know about teachers. We know how much they are paid on average, we know roughly how much professional training they do each year, we know how many hours they spend teaching in class. But how much do we know about what teachers do in these classrooms that helps students do well academically, socially and emotionally?

Questions are being asked. Since the Programme for International Student Assessment (PISA) of 15-year-olds was carried out in 2015, expenditure on schooling has climbed steadily. And yet, three years later, PISA shows that students' performance scores in reading, mathematics and science in the Western world have flat-lined.

This discouraging scenario becomes all the more so as we grapple with the effects of the COVID-19 pandemic. As public revenues decline, education spending is in danger of flat-lining as well, or, even dipping. Should budgetary choices have to be made, it is necessary to know what helps students learn and what does not.

This report sets out some insights on what works best. They have been gleaned from the 2018 PISA, together with the 2018 Teaching and Learning International Survey (TALIS) of teachers and school leaders on school environments. AI-powered supervised statistical learning was used to finely comb through these combined datasets in search of clues – significant links between certain teacher characteristics and conditions in schools and classrooms, and students' PISA performance and responses about how they perceive everything from their teachers, parents, and schoolmates to their academic capabilities and future plans. These findings, fleshed out with existing research literature, confirm something we have suspected all along: teachers and schools make an important difference to how a student performs and feels. More specifically, it is the time teachers spend actually teaching in class, not disciplining or taking care of administrative work, and the hours they spend marking and correcting work, and going over this feedback with their students that links to how well students do academically, and how motivated and optimistic they are about their learning and prospects.

But the question is not merely what teachers and schools can do to help students flourish, but what they can do to help the greatest number of students flourish. To varying degrees, our education systems struggle with inequalities that influence the academic outcome and, ultimately, the life trajectory, of each student. Some of those inequalities are rooted in gender difference. This report reveals that 15-year-old boys perform markedly worse on PISA than girls when there are disciplinary problems in the classroom. But it also shows that boys respond well to regular testing and parental involvement in school life. They also do better than girls academically when their teachers undergo regular performance reviews.

Socio-economic inequality drives another wedge between students, one that the COVID-19 pandemic widens. Stop-and-start schooling has upended most students' lives in 2020 and perhaps beyond. Temporary school closures have made learning especially challenging for vulnerable students who have poor access to the Internet, computers, or a quiet place in which to do their schoolwork. Some students may quit school altogether.

All this makes it crucial that schools and education systems get it right. This report finds that classroom teaching, and teacher feedback and assessments go a long way to helping all students, regardless of who they are or what their background is. So do classrooms that mix socio-economically advantaged and disadvantaged students because the presence of the former elevates the achievement, sense of capability and aspirations of the latter. And there is one last thing: students do better when their parents and guardians, and local communities involve themselves in school activities.

Arranging teachers' workload and schedules so they can devote more time to actual teaching, and putting students of mixed socio-economic backgrounds together are, quite frankly, meat-and-potatoes education policies. They lack the sizzle of "anytime, anywhere" collaborative online learning and other tech tools we believed would kick school systems into 21st-century gear. But the good news for what certainly will be budget-strapped governments is that these and other recommendations in this report are not especially hard or expensive to implement.

The recommendations are also remarkably prescient. Though these TALIS and PISA surveys were carried out long before COVID-19 hit, their data tell us exactly what we have been learning about remote learning during this pandemic: teachers and schools really do matter for students. Most parents have newfound respect for what teachers do in their classrooms. And, even if they like the flexibility, most students realise they miss interacting with their classmates and learning in school. The report's conclusions about the importance of teachers and schools strangely anticipated ours once we had to cope without them.

Disrupted schooling leaves today's students less prepared for finding good work in what is increasingly a knowledge economy. To catch them up on compromised cognitive and socio-emotional skills, we must mobilise everything we know about how students learn best. This report tells us that what works best for every student has been there all along – right under our nose.

1 Overview: What the TALIS-PISA link insights imply for policy and future research

This OECD report on the results of the TALIS-PISA link 2018 focuses on the many ways teachers and schools matter for student achievement and social-emotional development. This chapter provides an overview of the datasets upon which this report is based, as well as the statistical methods – including a machine learning technique – used to analyse them. It also provides an overview of the report's main findings, followed by several directions for education policy. Finally, it offers recommendations for improving the current survey design in order to better examine and understand the connection between teaching and learning in the future.

Introduction

Positive, High-achieving Students? What Schools and Teachers Can Do acknowledges the multiplicity of teaching and schooling goals. Not only do teachers need to provide students with cognitive skills so they can find employment and thrive in life but they are also expected to efficiently select and sort students based on their abilities and interest, promote equality of opportunity, and provide students with the skills to become active citizens (Van De Werfhorst and Mijs, 2010[1]). This reports sheds light on the many ways teachers contribute to these goals and, ultimately, the development of children and young people.

The pressure put on teachers has only become greater since the outbreak of the COVID-19 pandemic. With school closures, the rapid expansion of distant and blended learning, and physical distancing needs at school, teachers have had to rethink the way they teach, organise their time, and interact with students, parents and colleagues. More than ever, education systems are revisiting the way they train their teachers and support them in changing learning conditions. The findings of this report can help educational stakeholders review their practices as it provides ample illustration of the many ways in which teachers are effective for student achievement and socio-emotional development and for ensuring that all students thrive, no matter what their personal characteristics and circumstances.

Identifying the teacher and school factors that help younger generations to succeed and thrive later in life has long been a main interest for education policy. Past education research has shown that how teachers, school leaders and schools shape the quality of instruction and students' environment impacts strongly on student academic and social-emotional development (Barber and Mourshed, 2009[2]; Darling-Hammond, 2017[3]; OECD, 2018[4]).

Past studies have found that teacher quality is the most important school-related predictor of student achievement (Hattie, 2009[5]; Rice, 2003[6]; Seidel and Shavelson, 2007[7]; Wayne and Youngs, 2003[8]). "Teacher effect"[1] (or a teacher's value-added) accounts for significant variation in student achievement (Chetty, Friedman and Rockoff, 2014[9]; Chetty, Friedman and Rockoff, 2014[10]; Hanushek and Rivkin, 2010[11]; Jackson, Rockoff and Staiger, 2014[12]; Kane and Staiger, 2008[13]; Rivkin, Hanushek and Kain, 2005[14]; Rockoff, 2004[15]). The estimated magnitude of teacher effect is relatively large compared to the effects of other school factors (Jackson, Rockoff and Staiger, 2014[12]).

There is also evidence that teachers matter for student social-emotional skills. Studies on teacher value-added have found that, as with test scores, teachers vary considerably in their ability to impact students' social and emotional development, including a variety of observed school behaviours (Gershenson, 2016[16]; Jackson, 2018[17]; Jennings and DiPrete, 2010[18]; Koedel, 2008[19]; Kraft, 2019[20]; Ladd and Sorensen, 2017[21]; Ruzek et al., 2015[22]).

Building on this evidence, this report aims to identify the specific characteristics and actions of teachers and schools that matter for student achievement and social-emotional development. While doing so, it acknowledges the growing body of research literature suggesting that teacher effects are multidimensional. First, teacher effects can differ depending on the student outcome of interest. Several studies found that teachers who make a difference for students' academic performances are not necessarily equally good at contributing to students' social and emotional development (Gershenson, 2016[16]; Jackson, 2018[17]; Jennings and DiPrete, 2010[18]; Kraft, 2019[20]; Ruzek et al., 2015[22]). Second, teacher effects are susceptible to variation, depending on student background. For instance, teachers may interact differently with boys and girls, or with socially disadvantaged and socially advantaged students, depending on their beliefs about their interests and abilities (Hadjar et al., 2014[23]). In addition, certain teaching strategies may be more effective among certain groups of students than among others (Le Donné, Fraser and Bousquet, 2016[24]; OECD, 2015[25]). Consequently, this report aims to address the following questions:

- Chapter 2: What do schools and teachers do that matters most for student achievement?

- Chapter 3: What do schools and teachers do that matters most for student social and emotional development?
- Chapter 4: Does what teachers and schools do matter differently depending on student achievement and gender?

Chapter 1 summarises the results highlighted across the report and tackles the cross-cutting question of the multidimensional nature of teacher and school practices and policies – do teacher and school factors matter equally for all student outcomes and for all students? Or, do some factors matter more (or differently) for certain student outcomes and groups of students? Chapter 1 presents policy implications based on the main findings of the report. It also highlights a number of unresolved questions and offers recommendations for future data collection and avenues for further research.

The TALIS-PISA link data

To address these important research and policy questions, this report draws entirely upon combined TALIS 2018 and PISA 2018 data.[2] Since the first cycle of TALIS in 2008, the OECD has striven to describe teaching and learning environments from the point of view of teachers and school leaders. Likewise, since the first cycle of PISA in 2000, the OECD has endeavoured to measure student achievement and social-emotional skills, and identify school practices and policies that are associated with the performance and equity of education systems. With the TALIS-PISA link option, the OECD and the participating countries and economies have put together these two parts of the puzzle – TALIS and PISA – to obtain as complete a view as possible of what happens in today's schools.

TALIS 2018 provides substantial data from nationally representative samples of teachers on the background, beliefs and practices of lower secondary teachers and the principals of their schools (OECD, 2019[26]). It is the largest international survey that focuses on the working conditions of teachers and school leaders, and the learning environment in their schools. TALIS aims to provide valid, timely and comparable information to help countries review and define policies for a high quality teaching profession (Ainley and Carstens, 2018[27]; OECD, 2019[26]). It is an opportunity for teachers and school leaders to provide input into educational policy analysis and development in key areas of their work (OECD, 2020[28]; OECD, 2019[29]).

PISA 2018 delivers insights into the family and school background, and cognitive and social-emotional skills of 15-year-old students. It assesses to what extent children near the end of compulsory education have acquired the knowledge and skills needed in modern societies (OECD, 2019[30]; OECD, 2019[31]; OECD, 2019[32]). With regard to student achievement, this report uses student performances in all three subject domains traditionally assessed by PISA (reading, mathematics and science) with more weight given to reading since it was the focus domain of PISA 2018[3] (Chapters 2 and 4). In addition, the report analyses a broad range of social-emotional development indicators (Chapter 3) and examines four that vary more across schools in greater depth. These are: students' perceptions of their classroom climate, their teacher's enthusiasm for teaching, the difficulty of the PISA test, and students' expectations of completing at least a tertiary degree. Student background information – particularly gender, socio-economic and cultural status, and migration background – is used in the analysis to neutralise potential mediating effects of student characteristics (in all chapters) so as not to skew relations of interest. In Chapter 4, student background information is used to break down the student population into sub-groups – notably girls and boys – and estimate gender gaps in student achievement within schools.

TALIS offers data on several dimensions of teachers' and principals' work, as well as school characteristics and practices, which contextualise and frame students' cognitive performance and social-emotional skills as measured in PISA. Linking PISA with TALIS creates a rich dataset enabling student, teacher, principal and school data across countries to be connected.

A great advantage of this dataset is its international component. Most of the studies surrounding teacher and school effects are conducted in one specific country with little possibility of exploring whether the findings also hold for other national contexts. In TALIS 2018, participating countries and economies had the option of administering TALIS questionnaires to a PISA 2018 sub-sample of schools with the purpose of linking schools', principals', teachers' and students' data. Nine diverse countries and economies from four continents took part: Australia, Ciudad Autónoma de Buenos Aires (hereafter CABA [Argentina]), Colombia, the Czech Republic, Denmark, Georgia, Malta,[4] Turkey and Viet Nam[5] (Box 1.1).

Box 1.1. The TALIS-PISA link 2018

The main features of the TALIS-PISA link 2018 are as follows:

- Participating countries/economies: Australia, CABA (Argentina), Colombia, the Czech Republic, Denmark, Georgia, Malta, Turkey and Viet Nam.
- Representative samples of schools and 15-year-old students within schools with a target nominal sample size of 150 schools per country and 35 students in each school. In each PISA-participating school, the school principal and a random sample of 20 teachers teaching 15 year-old students were selected.
- TALIS technical standards on response rates: the minimum response rate is 75% of sampled eligible and non-excluded schools and 75% of all sampled teachers across all participating schools.
- PISA technical standards on response rates: the minimum school response rate is 85% of sampled eligible and non-excluded schools and the minimum student response rate is 80% of all sampled students across all participating schools.
- TALIS teacher and school principal questionnaires were administered online and on paper. Some TALIS questions were adapted to refer to 15-year-old students.
- PISA questionnaires, including, in particular, the student questionnaire, as well as student assessments in reading (the main domain in 2018), mathematics and science were administered in computer- and paper-based mode.
- Data collection window: the TALIS questionnaires were administered in parallel with or shortly after the administration of the PISA instruments during the same school year.

Sources: OECD (2020[33]), *PISA 2018 Technical Report*, https://www.oecd.org/pisa/data/pisa2018technicalreport; OECD (2019[26]), *TALIS 2018 Technical Report*, http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf.

The specific survey design of the TALIS-PISA link data has important implications for the interpretation of the results presented in this report. First, the link between the TALIS and PISA surveys operates at the school level and not at the class level. In other words, the data do not allow matching between a teacher and her or his students; rather, the data only permit matching between a sample of teachers teaching 15-year-old students in a school and a sample of 15-year-old students of that same school. While analysing the data, three types of links can be established between the TALIS and PISA data:

1. Individual student data can be merged with TALIS data from teachers and school principals aggregated at the school level.
2. Individual teacher and principal data can be merged with PISA data from students aggregated at the school level.
3. TALIS data aggregated at the school level can be merged with PISA data aggregated at the school level.

In this report, most of the analyses are conducted by connecting individual student data with data from teachers averaged at the school level and with school principal data. What is measured by TALIS data at the school level relates to a school's overall context and needs to be interpreted accordingly. The analyses presented in this report generally link the characteristics and practices of the average teacher in the school with the cognitive or social-emotional skills of individual students in the school. Therefore, the reported analyses are likely to be conservative to the extent that an association between an average teacher and individual student outcomes might not be significant due to the aggregation of different teachers' data at the school level even though, in reality, an individual teacher might contribute to an individual student's outcomes.

Another limitation of the study design is that both the TALIS and the PISA studies are cross-sectional, i.e. they measure student, teacher, principal and school characteristics in many countries, but at a single date. This survey design prevents causal interpretation of the reported analyses. The associations between student outcomes, and teacher and school factors highlighted in this report need to be cautiously interpreted and do not allow drawing any conclusions about teacher or school effectiveness.

Finally, it must be noted that this report draws on data collected in 2018, i.e. before the outbreak of the COVID-19 pandemic. Therefore this report cannot be informative about the teacher and school factors that have become more relevant for student cognitive and social-emotional development following crisis-induced changes, such as school closures, the rapid expansion of distant learning and teaching, or the imposition of physical distancing in the schools. Nevertheless, this should not unduly lessen the report's reach as the analyses aspire to some form of generality. They aim to identify teacher and school factors that generally contribute to student fulfilment. This quest goes beyond the contingencies of the current context. Some kind of generality can be claimed, for example, whenever a teacher factor is found to be consistently associated with a student outcome across several countries and/or across approaches. In particular, some general conclusions are formulated when a relation is found to hold on average across the participating countries/economies and at least (and sometimes only) a couple of countries. In any case, this report is conceived as a generator of hypotheses that warrant further examination in future studies.

Use of advanced statistical methods, including a machine learning technique

To make the most of the TALIS-PISA link data, this report takes a mixed data and theory-driven approach. The TALIS 2018 conceptual framework (Ainley and Carstens, 2018[27]) and previous research findings are carefully considered to inform the analyses and interpret, validate, or contextualise the findings. Each of the following chapters builds on the research literature and presents a theoretical framework that underpins the analyses.

The approach is also strongly data driven insofar as the identification of the teacher and school characteristics and practices that matter for student outcomes is centred around a machine learning technique called "lasso" (see Box 2.1 in Chapter 2 for a presentation). Machine learning techniques have rapidly developed with the recent explosion of "Big Data" (Hastie, Tibshirani and Friedman, 2017[34]). They are an attractive tool for analysing data patterns emerging from the many variables collected through the TALIS questionnaires and the many student outcomes measured by the PISA assessment and questionnaire. In particular, lasso can select variables that are highly correlated with the outcome variable even when the number of potential variables is high relative to the number of observations. Lasso regressions are performed in each chapter as a compass guiding the selection of key teacher and school factors related to student achievement, social-emotional skills and gaps in student performance within schools. This OECD report is the first of its kind to apply this machine learning technique to large-scale international education data.

Variance decomposition techniques are applied to measure the share of variance in student outcomes explained by each of the teacher and school factors considered (those highlighted by lasso and by past research). Standard statistical methods are then used to determine the significance and sense of the relationships between teacher and school factors, and student outcomes. These standard linear and logistic regressions are applied country by country and on average across countries/economies. Finally, quantile regressions are used in Chapter 4 to delve deeper into whether some of the key identified associations vary in some way for students with different competence levels. The interpretation of the findings focuses on the significance and the sign of the regression coefficients. While the size of coefficients are displayed in figures and annex tables, they are not commented in the report for the sake of brevity and due to reservations regarding over-reliance on them.

By investing in several advanced statistical methods, including a machine learning technique, this report seeks to break new ground in extracting maximum relevant information from a complex dataset. More information on the statistical analyses conducted in this report can be found in Annex B.

Overview of the main findings

Schools make a difference in how students perform, behave and feel

Schools and their teachers can actually make a difference not only for student cognitive performances but also for a wide range of students' social-emotional outcomes, including student school behaviour, interest in school, self-concept and education aspirations. Differences in school average performances represent a bit more than 30% of the total variance in student performances, irrespective of the subject domain, on average across participating countries and economies. Schools also differ significantly in the way students perceive their classroom climate and their teachers' enthusiasm for teaching – around 15% of the total variance in these outcomes lies between schools on average across countries. There are also significant differences between schools in students' perceptions of the PISA test difficulty and in students' expectations of completing at least a tertiary degree (10% or more of the total variance on average across countries). School differences result from two phenomena that can mutually reinforce or mitigate each other: prior school segregation of students by their academic, socio-demographic or social-behaviour characteristics, and variations in schools' ability to improve student outcomes. Unfortunately, in the absence of a longitudinal design, the TALIS-PISA link data do not allow disentangling both phenomena and quantifying how much of the school differences each of these phenomena account for.

One might think that schools are more likely to differ in student achievement than in student social-emotional development as schools are more traditionally focused on the former. Yet, research conducted in the US context suggests that teacher effects on certain student attitudes and behaviours are similar in magnitude or even greater than teacher effects on student test scores (Blazar and Kraft, 2017[35]). Little is known about the way teacher effects aggregate at the school level, i.e. whether they tend to cumulate or cancel each other out within schools, and whether this varies depending on the outcome of interest. The findings from this report may suggest that teacher effects on student social-emotional skills do not cumulate as much as teacher effects on student achievement do. However, there are other more plausible reasons why schools appear to differ more with respect to student achievement than student social-emotional outcomes (based on PISA data). These have to do with the fact that student achievement is better measured than student social-emotional outcomes in PISA. While student performances in each domain are estimated on at least half an hour of test, indicators of student social-emotional development are measured based on one question each.[6] In addition, students' answers to questions about their attitudes, beliefs and self-concept are more likely to suffer from reference bias than students' answers to the PISA cognitive test. For example, PISA 2018 questionnaire asks students to rate themselves on statements such as: "I am a good reader"; "I am able to understand difficult texts": or "I normally feel happy"; "I normally feel proud", etc. In answering these questions, students must interpret the definitions of reading proficiency or happiness, for example, which likely involves comparing themselves to other people, students from their class and school, in particular. It is likely that students from different schools have different standards or reference points when answering these questions. In such cases, the differences between schools in students' social-emotional outcomes might well be underestimated.

In any case, the analyses presented in this report can only explain the share of variance in student outcomes that lies between schools as teacher and school factors are all introduced at the school level. TALIS-PISA link data are more likely to provide insights related to student performance for the Czech Republic and Turkey where differences in school average performances represent about half of the total variance in student achievement as opposed to countries, including Australia, Denmark and Malta, where 25% or less of the total variation in student outcomes lies between schools. Cross-country patterns also emerge regarding school differences in student social-emotional development. The Czech Republic and Georgia are consistently among the countries with the largest school differences in student social-emotional outcomes. At the other end of the spectrum, Colombia and Denmark are consistently among the countries showing the smallest school differences (yet still at around 10% of the total variance or above).

The structure of these education systems, in particular the number of education programmes available to 15-year-olds and students' age at first selection, might explain part of the differences resulting from school segregation. For example, the large differences observed in the Czech Republic might be explained by the fact that 15-year-old students can be enrolled in five different education programmes and they are first selected into these tracks at the age of 11. On the other hand, in Denmark, where small differences are observed between schools, the age at first selection is 16 and all 15-year-old students are enrolled in the same education programme – see OECD (2020[36]), Figure V.3.2.

Teacher and school factors selected by lasso, a supervised statistical learning method

Lasso regressions are performed in Chapters 2 to 4 as a compass guiding the selection of key teacher and school factors related to student achievement, social-emotional skills and gaps in student performance within schools. Four teacher and school factors are consistently selected by lasso as key predictors of both student achievement and student social-emotional development: teachers' classroom practices, classroom characteristics, school culture and school leadership. Two teacher and school factors are uniquely selected by lasso as key predictors of student achievement: teachers' use of working time and teachers' well-being and job satisfaction. This selection of key teacher and school factors is established on the overall population of students, teachers and principals surveyed within all countries and economies participating in the TALIS-PISA link. Variance decomposition analysis indicates that these factors selected by lasso explain at least 20% of the variation between schools in student achievement, and at least 10% of the variation in student social-emotional development.

The four teacher and school dimensions highlighted by lasso as important for student achievement across subjects and student social-emotional development represent teacher and school dimensions with direct and indirect effects on student outcomes. The fact that only four out of the 18 dimensions are found to matter across different student outcomes suggests that the relationships between the teacher and school factors are often characterised by reciprocity and interconnectedness. Indeed, most teacher and school factors such as teacher training or teacher collaboration that are not selected by lasso certainly matter for student outcomes, if only indirectly given the reciprocal and interconnected nature of the relationships between the different dimensions.

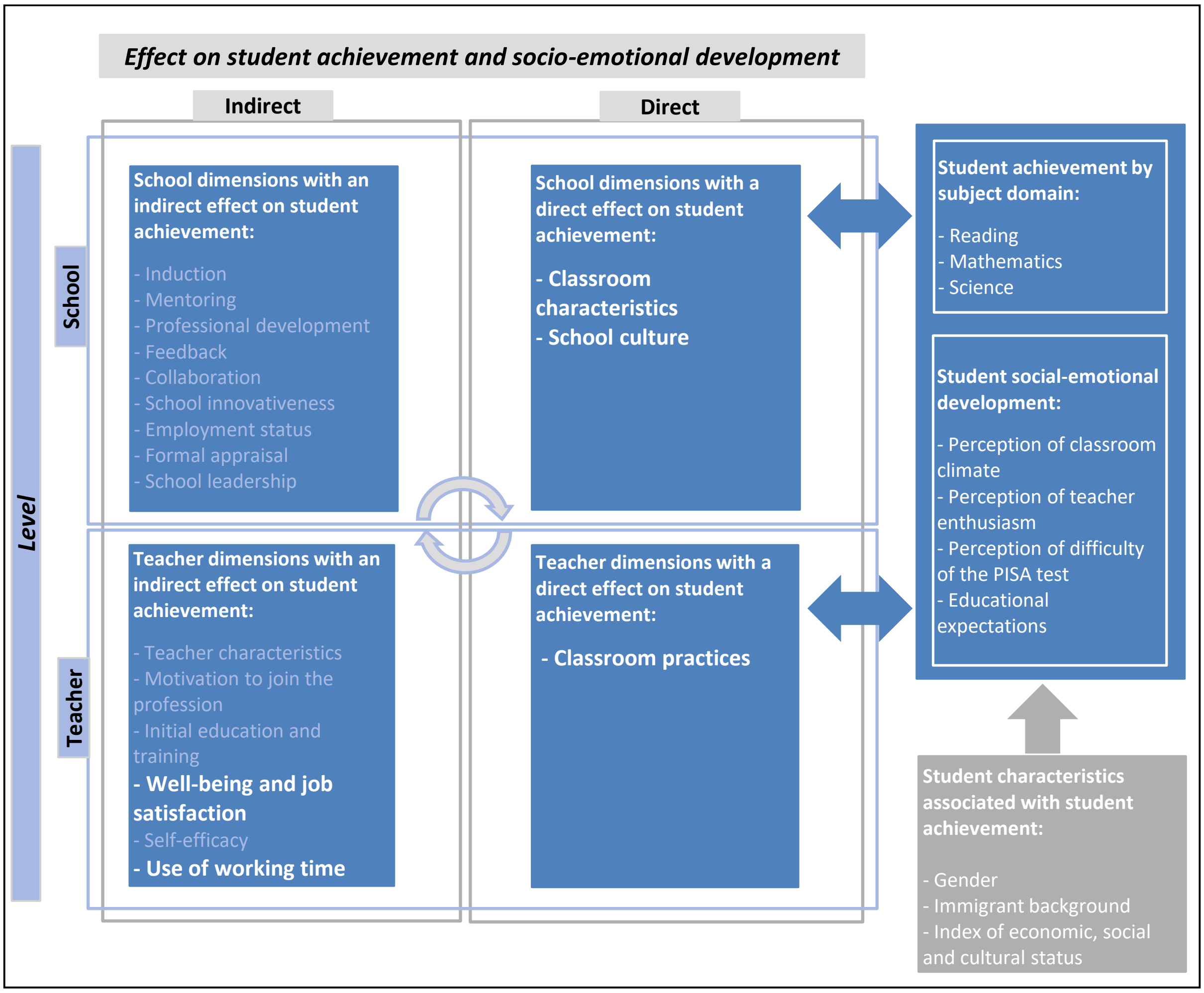
The teacher and school factors selected by lasso are also highlighted by the research literature as important elements in relation to student achievement and social-emotional development. The relationship between teachers' classroom practices and student achievement as well as student attitudes and behaviours is well established in past research (Blazar and Kraft, 2017[35]; Hattie, 2009[5]; Le Donné, Fraser and Bousquet, 2016[24]; Muijs et al., 2014[37]). The importance of classroom characteristics, in particular the classroom's student composition in terms of socio-economic background and ability level echoes what education research finds in relation to peer effects (Carrell and Hoekstra, 2010[38]; Gaviria and Raphael, 2001[39]; Sacerdote, 2011[40]). Similarly, past research also established the link between school-level factors such as school culture and school leadership, and student outcomes, especially academic achievement (Ainley and Carstens, 2018[27]; Chapman et al., 2015[41]; Hallinger, 2018[42]; Hallinger, 2015[43]).

Teacher and school factors that matter for student academic success and social-emotional development

Standard regressions allow a more detailed focus on specific characteristics and practices of teachers and schools that matter for student achievement and social-emotional development. These factors are: teachers' classroom practices, teachers' use of working time, teachers' well-being and job satisfaction, classroom characteristics and school culture (Figure 1.1). This section describes the findings of Chapters 2 and 3 that hold on average across the TALIS-PISA link countries and economies, and for at least two countries.

What teachers do in the classrooms matters for both student achievement and social-emotional skills, particularly student behaviour in class. Students tend to perform better on average the more class time teachers spend on actual teaching and learning. This finding suggests that students' opportunity to learn is important for student achievement, and this is borne out by the research (Schmidt, Zoido and Cogan, 2014[44]). Clearly, students' opportunities to learn are closely linked to the amount of time allocated to academic instruction (Muijs et al., 2014[37]) and the link between the quantity and pacing of instruction, and student achievement is also consistently confirmed by past research (Muijs et al., 2014[37]). The amount of time students are actively engaged in learning during a lesson is, in turn, highly related to the classroom's disciplinary environment and the teacher's classroom management practices and skills (Muijs et al., 2014[37]). When teachers do not actually teach and students do not learn in class, it is usually either due to disciplinary issues or administrative tasks. As previous TALIS findings show, experienced teachers spend more time on actual teaching and learning partly because they teach in less challenging classrooms (OECD, 2019[29]). The positive association between the share of class time spent on actual teaching and learning – as opposed to administrative and classroom management tasks – and student achievement may also point to a reverse causal relationship. More disruptive classrooms are more likely to have lower-achieving students, which, in turn, leads to more time spent on other tasks such as keeping order and administrative tasks. TALIS-PISA link results suggest that opportunities to learn are more closely linked to student performance in mathematics than in reading and science.

Figure 1.1. Teacher and school factors that matter for both student academic success and social-emotional development



Teacher-student interactions are also related to student behaviour in class. Results show that the more teachers report nurturing good relationships with students, the more students perceive them as enjoying teaching and the better it is for classroom disciplinary climate. In addition, there seems to be clear school patterns of disciplinary climate. There is a high degree of alignment in different stakeholders' perceptions of school climate[7] and consistency in the measures taken by school leaders and school teachers to address disciplinary issues. First, the more teachers report disciplinary issues, the more students report the same. Likewise, the less disciplined the classroom climate is, as reported by students, the more frequently principals report collaborating with teachers to solve classroom discipline problems, and the more time teachers report spending on communication and co-operation with parents. This does not necessarily mean that principals and teachers' remedial measures are detrimental to student behaviour in class. Rather, this likely means that teachers spend more time on these kinds of activities when they feel students need it.

The findings of this report suggest that such vicious circles can be inverted by establishing positive relationships between teachers and students, and promoting a favourable climate for teachers' work well-being. Results show that students tend to find their teachers more interested in their teaching when teachers report lower levels of work-related stress on average in the school. Teachers' job satisfaction also matters for student performance. The more satisfied teachers are with their work environment, the better students tend to perform in school.[8] This finding suggests that teachers' satisfaction with their work environment can play a role in teachers' attitudes, efforts and commitment, which, in turn, can eventually lead to better performance. But it can also signal the presence of self-enforcing dynamics. Teachers may be particularly satisfied when they work in schools attended by high-achieving students. In turn, these teachers might be particularly committed to helping their academically gifted students progress further. In addition, results signal the presence of differential effects across subjects as teachers' satisfaction with their work environment seems to be more closely related to student performance in science than in reading and mathematics. While this may be explained by the fact that, unlike reading, students mainly acquire their knowledge in science at school, it can also point to the fact that certain requirements regarding the work environment, such as a well-equipped school laboratory, are critical for science teachers to do their work properly.

Teachers and principals are not the sole stakeholders that matter for students' academic and social-emotional skills. Parents and the local community also play an important role in student achievement, especially when it comes to their involvement in school-related activities.[9] Nevertheless, the association between parental involvement and students' academic performance no longer holds after classmates' characteristics are taken into account. Indeed, classmates seem to matter a great deal for student performance and student self-concept. First, the concentration of socio-economic disadvantage among classmates seems to matter for both student achievement and student educational aspirations. As the average concentration of students from socio-economically disadvantaged homes in the classrooms increases, students tend to perform worse academically and be less likely to aspire to tertiary education studies. Since these findings hold while accounting for students' own socio-economic background, it suggests the presence of peer effects. Such concentration of disadvantages may affect the student's cognitive development due to fewer available material learning resources at the school and altered teaching strategies. Students might also be more easily influenced by the lower aspirations of their socially disadvantaged peers or by the shortage of highly educated adults as role models among other peers' parents and, by extension, the local community. Yet, consistent with past PISA-based findings (Raitano and Vona, 2013[45]; Rangvid, 2007[46]), high-performing students still tend to be less affected than their low-achieving peers by the composition of their classes. This suggests that addressing socio-economic segregation of schools may be beneficial for both increasing student performance at the country level as well as improving equity in educational achievement and opportunities.

In addition, the TALIS-PISA link data point to some encouraging findings. First, the presence of academically gifted students in the classroom seems to be beneficial to students' self-concept in the short term, and to their PISA scores. Indeed, the greater the number of academically gifted students enrolled in the classroom, the more students feel able to succeed in the PISA test. They also end up performing better on average. These findings may signal the presence of academic segregation as high achievers tend to be concentrated in certain schools in most education systems but it can also point to the presence of peer effects. Students' performance can be positively affected by classmates with higher innate ability through an increase in motivation, competition and career aspirations (OECD, 2019[31]; Sacerdote, 2011[40]).

Second, results show that the larger the peer group, the better it is for student self-concept, both in the short term (perception of improved performance on the PISA test) and the longer term (expectation of completing a tertiary degree). While causality cannot be determined, this might be due to the fact that the larger the class is, the more likely the student is surrounded by peers with various family backgrounds, attitudes, behaviours and aspirations. This creates a richer and more favourable context for the student to grow socially and emotionally. Yet, most likely, the presence of a selection bias is also partly at play. A school's average class size partly captures unobserved characteristics about school intake such as students' social-emotional skills. It is common practice for schools and school systems to implement specific grouping strategies and adjust class sizes to the needs of their students. A previous PISA report found that there was a significant difference of more than three students per class between socio-economically advantaged and disadvantaged schools on average across the OECD, and that such a positive and significant difference was found in 39 education systems (OECD, 2018, p. 89[4]).

Finally, although causality cannot be established, a few teacher characteristics and practices have been highlighted as being especially important for student social-emotional development: the level and quality of teacher pre-service training and the way teachers use their working time. First, results show that the greater the number of teachers who hold a master's or a doctorate degree within the school, and the more they felt prepared to teach their subject after their initial training, the more likely students are to expect to complete at least a tertiary degree. This might be because these teachers are better informed on the existing tracks to complete a tertiary degree, or because they represent higher educational aspirations for their students and influence them. Second, the more time teachers spend on extracurricular activities, the more students report that the classroom is disciplined, that the teacher is interested and motivated to teach, and that they expect to complete at least a tertiary degree. Past research found that participation in extracurricular activities can help students develop social-emotional skills such as persistence, teamwork or a stronger sense of belonging at school (OECD, 2020[36]) though recent research (Farb and Matjasko, 2012[47]) did not seem to have studied the particular effect of teachers' participation in these activities on student outcomes. In addition, the more time teachers spend on marking and correcting student work, the stronger students' self-concept of doing well in the PISA test and, eventually, pursuing tertiary studies. While the findings are slightly less robust on student achievement, they also suggest that the more hours teachers spend on marking and correcting student school work, the better it is for student performance. These findings suggest that spending quality time with students outside of the usual lessons and devoting time to assess student work and indicate progress to students are valuable ways to support student growth.

Teacher and school factors that matter more specifically for equity in performance

The aim of schools and education systems is for all students to be able to thrive, irrespective of their personal characteristics. This means that teaching practices and teacher-related policies benefit as many students as possible without being detrimental to some of the students – whether the most fragile or most academically gifted – and in a way that achieves the collective optimum. Chapter 4 tackles two distinct policy and research questions on equity in educational achievement:

- Do teacher and school factors matter equally for students with different academic performances?
- How can schools mitigate gender gaps in student performance?

The former section highlighted teacher and school factors found to be important for student performances on average. Could it be that these factors matter for the average student because they actually matter mainly for high-achieving students but not necessarily for low-achieving students? Results presented in Chapter 4 suggest that the answer to this question is no. On average across the countries and economies participating in the TALIS-PISA link, teacher and school factors that matter for average students' performances in reading, mathematics and science also tend to matter for both low and high-achieving students. These factors include the share of class time that teachers report spending on actual teaching and learning, the working hours teachers report devoting to correcting and marking their student work, teachers' satisfaction with the work environment as well as the concentration of academically gifted students in the classroom. In addition, parental and community involvement in school-related activities also matters for all students, regardless of their performance level. This suggests that all these practices could be leveraged to equally support student academic growth, regardless of their initial competence level.

One teacher factor matters specifically for low-achieving students – workload-induced stress. The share of teachers reporting workload as an important source of stress at school is positively associated with better performances for low-achieving students in two out of the three subject domains (reading and mathematics). This result aligns with past research that found that stressful working environments and challenging working conditions can affect the practices of teachers and principals, their work motivation, and even student achievement (Viac and Fraser, 2020[48]). Yet, research has associated high levels of stress with deteriorated teacher outcomes such as lower self-efficacy for teaching and lower job satisfaction (Collie, Shapka and Perry, 2012[49]). Findings from this report suggest that being collectively affected by increased workload is not the same thing as being the only teacher in the school feeling stressed due to workload. This might point, generally, to teachers' commitment to their work, which might be particularly beneficial to low-achieving students. Yet, stress can also develop into burnout. Thus, workload-induced stress also signals the need for support for committed and dedicated teachers who feel overwhelmed with the workload.

Equal opportunities for all students also means that differences in students' outcomes are not driven by individual characteristics and circumstances such as gender, socio-economic status or immigration background, everything else being equal. Chapter 4 focuses on gender gaps in student achievement and teacher and school factors that could mitigate them. Two main findings can be drawn from this chapter.

First, boys seem to be more disturbed than girls by classroom disciplinary problems and school organisational issues. As the level of disciplinary issues perceived by the average school teacher increases, the difference in PISA reading scores between girls and boys within schools tends to increase further in favour of girls. And the more often school leaders observe instruction in classrooms and resolve problems with the lesson timetable in the school, the more girls outperform boys in reading. Boys might be more affected by deteriorated learning conditions, as they tend to be exposed to greater peer pressure than girls. They may also more likely be the students disturbing lessons. This echoes several research studies suggesting that, for many boys, it is not acceptable to be seen to be as interested in school work as girls (DiPrete and Buchmann, 2013[50]). Boys adopt attitudes that include a disregard for authority, academic work and formal achievement (Salisbury, Rees and Gorard, 1999[51]), which might express more acutely when there are practical and organisation disruptions in the school. Overall, boys seem to be less able than girls to stay focused on their schoolwork when disciplinary or practical issues occur in the school. However, the findings also suggest potential measures to support boys in closing the gap with girls: greater involvement and support from parents in school-related activities as well as positive relationships between teachers and students. This suggests that, at the age of 15, boys (more than girls) are in need of support from significant adults – their parents and their teachers – to self-regulate and be achievement-focused.

Second, boys are more likely to perform as well as (or even better) than girls in reading in schools where a culture of student assessment, teacher accountability and appraisal prevails. The more often teachers evaluate their students by administering their own assessment, the smaller the difference in reading performance between girls and boys, in favour of boys. This may suggest that boys benefit from more regular testing as this allows them to better self-regulate and focus on schoolwork. Yet, this might also be due to girls' increased anxiety about frequent testing and induced competition (Devine et al., 2012[52]; Gneezy, Niederle and Rustichini, 2003[53]; McLean and Anderson, 2009[54]; Niederle and Vesterlund, 2011[55]). In addition, the more school leaders ensure that teachers feel responsible for their students' learning outcomes and the more often teachers are formally appraised by external individuals and bodies, the better boys perform in reading compared to girls. Giving teachers opportunities to reflect on their teaching practice and finding ways to support low- and middle-achievers, among which boys are overrepresented in reading, could help overcome gender inequalities.

TALIS-PISA link data throw a light on whether students performed better or not when taught by teachers of their gender. Analysis reveals some evidence that a student's performance tends to increase as the share of teachers of their gender in the school increases. This relation is observed for all three subjects in three of the seven[10] countries/economies included in the analyses – CABA (Argentina), the Czech Republic and Turkey – although not on average across the participating countries and economies. These findings support the notion that, at least in certain education systems, teachers may have a bigger influence on students of their gender either by applying distinct teaching practices that suit same-gender students better or by addressing gender stereotypes through acting as role models for these students (Dee, 2005[56]).

What these findings imply for policy

This chapter points out several directions for education practices and policies that would improve student performance and well-being. It includes a particular focus on improving low-achieving boys’ reading and student social-emotional skills. These directions draw on and revisit a few policy pointers among the many teacher and school practices and policies put forward in past OECD reports on TALIS 2018 and PISA 2018 results (OECD, 2020[36]; OECD, 2020[28]).

Once again, it is important to note that, given the cross-sectional design of the survey, the causality and directionality of relationships identified in TALIS cannot be ascertained. Moreover, there is no one-size-fits-all approach to designing policies to improve teaching and learning. When choosing among different policy options, governments must take into consideration the context of their education systems as well as a broad range of evidence to underpin and substantiate policy development. Accordingly, the policy considerations that follow should be interpreted as the OECD’s suggestions based on analysis results. They should be considered in each national context according to country-specific challenges and constraints.

Make the most of teachers’ time

An important precondition for quality teaching practices is to make the most of classroom time. Teachers could analyse the way they typically use their class time and, possibly in consultation with the school leader and other colleagues, find ways to reduce the class time spent on administrative tasks and keeping order in the classroom. This might imply finding alternative ways of dealing with administrative tasks and novel ways to improve relationships with students who are more likely to disturb classes. In particular, teachers’ involvement in extracurricular activities with their students might be an effective way to improve the disciplinary climate. Time spent with students outside of the regular classroom hours can be particularly positive for establishing and nurturing good relationships with students, which, in turn, are conducive of good conditions for students’ cognitive and social-emotional development.

Teachers should also find ways to improve the disciplinary climate of the classroom during regular classes. They could physically arrange the classroom so that students can focus or interact better, monitor their class time closely to manage it better, manage conflicts and allow students to work together and become more autonomous in the learning process. They should find ways to reduce noise in the classroom. They could, for example, use explicit visual codes to build a classroom noise barometer or invite students to experience silence. In their quest for solutions, teachers would also benefit from support by their colleagues, school principal, mentors, inspectors and instructors. They could also participate in professional development activities, focusing on one or several of these classroom management techniques.

Teachers not only need to make the most of their class time, they also need to devote sufficient time to activities that maximise student learning. This report sheds particular light on the importance of marking and correcting student work (for both student achievement and student social-emotional development). It reminds teachers, teacher trainers, school leaders and policy makers of the benefits of regular assessment. It not only provides teachers with an up-to-date picture of student progress, enabling them to determine what subject content and pedagogical approaches to use but it also gives students regular feedback on their progress and ways to improve. Teachers’ time spent on summative assessment can be indicative of their engagement in formative assessment. Regular formative assessment creates opportunities for constructive feedback, which helps build a trust-based relationship between teachers and students. Teachers, school leaders and educational authorities could review teacher work schedule to ensure that teachers devote sufficient time to analysing student work and communicating progress to students. Yet, as workload is an important source of stress for teachers, and marking/correcting student work is one of the most time-consuming tasks, education stakeholders could explore ways to use technology in such a way that teachers spend less time on marking while remaining committed to the formative part of assessment.

While this report highlights some teacher activities that seem particularly beneficial to students' development – participating in extracurricular activities and marking/correcting student work – teachers' workload should not be extended to include more of these activities unless the time they spend on activities such as administrative work, which serve student development less, is reduced.

Promote an achievement-focused culture within schools

Schools and school systems could promote a culture focused on achievement not only to raise performance but also equity. First, regular use of both formative and summative assessments can be useful for identifying the learning content that students still need to acquire and the learning methods that they could still improve. Students who receive regular feedback from teachers may be better prepared to correct deficiencies in the learning process. Moreover, regular testing of students provide an incentive for students to improve their learning outcomes, especially when the students are boys and when teachers administer their own tests. Such regular tests and feedback can give students a sense of agency, and a belief that they can influence their own learning. It can also strengthen schools' and teachers' sense of responsibility.

Second, school culture should not only put emphasis on the achievement of students but also that of teachers. School leaders can play an active role in continuously fostering teachers' feeling of responsibility over student performance. In turn, this type of leadership and sense of accountability can be particularly helpful in closing the gap between girls and boys by stimulating boys' performances. Likewise, a more frequent appraisal of teachers can be leveraged to promote both efficiency and equity in educational achievement. Formal appraisal by external individuals and bodies, which is considered more objective and less judgemental than appraisal by the school management team or other colleagues who work in the same school, may be particularly effective. Making a judgement and providing feedback to teachers about their competencies and performance is a building block of effective schools and education systems. In its summative form, appraisal is a tool for quality assurance, ensuring that required standards are met or recommended practices followed. Yet, there are risks that the formative function of formal appraisal is too closely associated with high-stakes appraisal for accountability purposes. Teacher formal appraisal could be enhanced by central standards and criteria, an independent appraisal component external to the school, a more formal and standardised process, and possibilities for teachers to appeal where they are doubts about the fairness of the process (OECD, 2013, p. 232[57]). Appraisal can also take a more formative emphasis and provide an opportunity for teachers to reflect on their practice, strengths and weaknesses in order to identify areas for improvement and career growth (OECD, 2019[29]). PISA 2018 results also find that greater equity can be attained when there is good balance between school autonomy and more centralised accountability measures (OECD, 2020, p. 201[36]). Formative teacher appraisal benefits from a non-threatening appraisal context, a culture of ongoing observation, mutual appraisal and feedback within the school, clear individual and collective objectives, simple school-based appraisal tools and supportive school leadership (OECD, 2013, p. 233[57]). More policy directions on teacher appraisal can be found in the OECD Reviews of Evaluation and Assessment in Education – see, in particular, Chapter 5 of the report *Synergies for Better Leaning* (OECD, 2013[57]).

Optimise student grouping and teacher assignments

Student segregation and teacher sorting across schools have been widely documented in past OECD reports (OECD, 2019[29]; OECD, 2018[4]). Although this is an important component of inequalities in student achievement and, likely, of inequalities in student social-emotional development, the issue of school segregation is considered to be beyond the reach of this report. Rather, this report examines how inequalities in student outcomes relate with what happens in the schools. It suggests policy levers that can mitigate such disparities.

First, schools can optimise the way students are grouped within classes in a way that is more profitable to the most fragile students than it is detrimental to the strongest students. Schools should strive to spread out both students with disadvantaged socio-economic background and academically gifted students as equally as possible across classes. This helps address not only disparities in student achievement but gaps in student self-concept, which are both mutually reinforcing.

Second, schools should optimise the assignment of teachers to classes. Students benefit from teachers with varied backgrounds (especially in terms of gender) across subjects and/or across school semesters/years. Variety in teacher backgrounds would allow for a variety of teaching approaches, (biased) beliefs about student ability and interest, and role models for students. Teachers and school leaders might consider inviting inspiring adults with characteristics that are under-represented among the school teaching staff into the classroom. Such interventions, even short ones, have proven to be effective at closing gender gaps in educational achievement, and educational and occupational aspirations (Breda et al., 2020[58]; Carrell, Page and West, 2010[59]).

Be attentive to teacher job satisfaction and well-being

Teachers' job satisfaction, well-being and beliefs are inter-related and have an indirect effect on student outcomes. Job satisfaction, which is the sense of fulfilment and gratification that teachers get from working, has a positive impact on teachers, school culture and ultimately on students (Ainley and Carstens, 2018[27]). Moreover, job satisfaction also plays a key role in teachers' attitudes, efforts and confidence in their daily work with children (Caprara et al., 2003[60]; Klassen et al., 2013[61]; Tschannen-Moran and Hoy, 2001[62]). Thus, positive job satisfaction may also have an indirect effect on student achievement through enhanced commitment leading to better performance of teachers.

This report suggests that some aspects of teacher well-being and job satisfaction are related to student achievement – teacher satisfaction with their work environment and teacher workload-induced stress. In consultation with teachers, school leaders and educational authorities could review working conditions in order to identify the areas that need to be improved. Such efforts could be especially beneficial in the most disadvantaged schools, whose teachers are most likely to face unfavourable working conditions. The fact that teachers' satisfaction with work environment matters particularly for student achievement in science might point to the importance of the quality of school equipment and resources for practical experiments in science lessons. It could also point to the fact that teachers of scientific domains might be more sensitive to their working conditions as they tend to enjoy more or better job opportunities in their field of work than other subject teachers. Therefore, schools and educational authorities should ensure that teachers have the required infrastructure and materials to deliver their subject lessons.

Teachers must also carefully organise their work and monitor their workload to ensure that the number of hours they work as well as the intensity of their work activities do not exceed critical thresholds. Yet, teachers are not the sole person responsible for their workload. School leaders and educational authorities should also ensure that the demands put on teachers are manageable within the time they have to complete their tasks. They must find balance between teachers' workload and the level of stress it induces. While a controlled dose of stress can be a potent force and reflect a feeling of commitment and dedication, teachers should not go too far and let workload-induced stress develop into burnout. It should also be assessed whether the teacher seems to be the only one in their school who considers workload to be an important source of stress. This might be particularly problematic as it could signal weaknesses in the practices of individual teachers. If other teachers in the school also experience workload-induced stress, it might reflect several things: a common commitment and dedication to the work; a general malfunctioning in the school; or a highly competitive school environment. School leaders, teacher trainers and inspectors could be involved to identify problems and provide support to solve them. In those cases where workload-induced stress is considered excessive, school leaders and teachers should aim to reduce administrative work through the enhanced use of digital tools or by rethinking and eliminating certain tasks. Administrative tasks are identified as one of the main sources of work stress by both teachers and school leaders (OECD, 2020[28]).

Engage with parents to build a positive school culture

Teachers and school principals could count more often on parents to help them build a positive learning environment in their schools. The family-school partnership can be encouraged through home-based parental support. Parents can be encouraged to discuss education matters with their child, help with homework and supervise their child's progress through education. Schools can also invite parents to interact more often with the school staff by communicating with school personnel, and participating in school decision making and school activities (LaRocque, Kleiman and Darling, 2011[63]). Schools can develop well-designed communications strategies to encourage parents to support their child's academic achievement and participate in school-related activities, in particular those parents who may be perceived as estranged from the school system or distrustful of the school. Schools can communicate to parents that their involvement in school allows parents to have a first-hand understanding of the learning environment, learn how to navigate the education system, demonstrate to their child that education is important, and influence their child's behaviour by establishing consistent norms (Grolnick and Slowiaczek, 1994[64]). Making the most of parents' involvement can also be a way to increase the school's human or material resources, especially when these are limited. Education authorities could also consider providing teachers and school leaders with training focusing on when and how to encourage parents' involvement in order to create a favourable school environment for students (LaRocque, Kleiman and Darling, 2011[63]).

Some schools and educational authorities have conducted very insightful experiments and experiences in creating community hubs, which link the school and the community and include education, health and facilities for all members of the community. More information about how local communities and parents can be mobilised to support schools can be found in Schleicher's report (Schleicher, 2018[65]).

What these analyses imply for future research

Exploiting TALIS and PISA data from nine countries and economies, this report has been able to shed some light on significant relationships between teacher and school factors, and student achievement and social-emotional development. Interpretation of the findings presented in the report is necessarily on the cautious side, given the limitations of the survey design. For this reason, this report is designed to put forward hypotheses that should be examined further in future research, whether on a larger international scale to test the generalisability of findings or on a more local field with a more suitable design to obtain more subtle results and/or to get greater insights on the sense and causality of the observed relationships. Accordingly, this chapter concludes with some final considerations regarding suggestions for improvements of future study designs.

In this report, the TALIS-PISA link data are used to contribute to an important goal of education research: determining what teachers' and school leaders' characteristics and practices affect students' cognitive and social-emotional outcomes and how. The TALIS-PISA link data present notable limits in addressing this ambitious yet policy-relevant question of which two are of particular importance. First, the link between teachers and students is only established at the school and not at the classroom level. Second, the data are cross-sectional and consist of a snapshot at a given time of teachers' and school leaders' characteristics and practices, and students' academic and social-emotional skills. This is why, at least in part, the sense of the relationships observed cannot be clarified nor their potential causality link be established.

Two main improvements could be considered for future projects linking TALIS and PISA data to improve understanding of the teaching-learning nexus. The first consists of linking the data from teachers with data from not merely students from the teacher's school who participate in PISA but with data from their own actual students. A possibility would be to propose to countries/economies participating in PISA that they over-sample classes of students enrolled in PISA modal grades (grades attended by at least one-third of PISA-eligible students)[11] and administer TALIS questionnaires to those teachers teaching the main PISA subject domains to PISA-participating students, if not to all of their teachers. While identifying and surveying these target populations may put additional burden on schools in some education systems, such study design would allow for a very good match between the students assessed and surveyed through PISA and the teachers surveyed through TALIS in the PISA-participating schools. As for any other OECD study, collected data from teachers and students would be processed with all necessary protection measures to minimise the risk of re-identification to be used for research purposes and in consultation with key social partners.

A second promising direction for improvement would consist of collecting data from the same students and teachers over time rather than (repeated) cross-sectional data. While longitudinal studies come with some drawbacks – being expensive and challenging insofar as participants need to be enrolled over an extended period in the study, their results are also stronger than those of cross-sectional studies. Student learning and social-emotional skills could be measured at different points of time (at least two points). This would allow for measuring changes in student outcomes over time and relating these changes with teachers' characteristics and practices at a given point of time. If data from teachers are also collected at different time points (during the period they teach the same students for obvious practical and analytical reasons), changes in student outcomes could also be related with changes in teachers' attributes. In comparison to repeated cross-sectional studies (TALIS and PISA are periodic surveys with independent samples for every cycle), longitudinal analysis can be used to make more conclusive findings (both at the individual level and at the aggregate level) than cross-sectional analyses, which are essentially descriptive. For example, by looking at within-student correlations (e.g. the correlation between teachers' working hours spent on assessing students and a student's outcome), one can rule out that stable between-student factors (e.g. student gender, socio-economic status, migrant background) can fully explain the association at stake.[12] Likewise, by looking at within-teacher correlations (e.g. the correlation between a teacher's job satisfaction and student outcomes), one can rule out that stable between-teacher factors (e.g. gender, initial teacher education, employment status, school assignment) can fully explain the association in question. In addition, interested countries and economies could also decide to conduct an experiment between two data collection points to examine the impact of a randomised intervention. It could then be possible to conclude a causal relationship between a given teacher practice and student outcomes.

The Global Teaching InSights (OECD, 2020[66]) was the first OECD education project to combine both a classroom-level link and a longitudinal design to describe teaching practices in the classroom and analyse their relationships with student outcomes. It provided very interesting insights about the nexus between teaching and learning. In particular, the study found that social-emotional support and classroom management were significant predictors of student personal interest and self-efficacy towards mathematics in half of the participating countries/economies, even after accounting for students' prior mathematics performance and other background characteristics. However, the outreach of the analysis was most likely limited by the use of small and non-representative samples, among other reasons. Drawing lessons from the Global Teaching InSights and the TALIS-PISA link option, efforts should be continued in this direction to design more valid, reliable and analytically powerful studies linking teachers and their students.

References

Ainley, J. and R. Carstens (2018), "Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework", *OECD Education Working Papers*, No. 187, OECD Publishing, Paris, https://dx.doi.org/10.1787/799337c2-en. [27]

Avvisati, F., N. Le Donné and M. Paccagnella (2019), "A meeting report: cross-cultural comparability of questionnaire measures in large-scale international surveys", *Measurement Instruments for the Social Sciences*, Vol. 1/8, pp. 1-10, http://dx.doi.org/10.1186/s42409-019-0010-z. [70]

Barber, M. and M. Mourshed (2009), "Shaping the Future: How Good Education Systems Can Become Great in the Decade Ahead", in *Report on the International Education Roundtable, 7 July 2009, Singapore*. [2]

Blazar, D. and M. Kraft (2017), "Teacher and teaching effects on students' attitudes and behaviors", *Educational Evaluation and Policy Analysis*, Vol. 39/1, pp. 146-170, http://dx.doi.org/10.3102/0162373716670260. [35]

Breda, T. et al. (2020), "Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools", *PSE Working Papers*, No. 2018-06, HAL Sciences de l'Homme et de la Société, https://halshs.archives-ouvertes.fr/halshs-01713068v3. [58]

Caprara, G. et al. (2003), "Efficacy beliefs as determinants of teachers' job satisfaction", *Journal of Educational Psychology*, Vol. 95/4, pp. 821-832, http://dx.doi.org/10.1037/0022-0663.95.4.821. [60]

Carrell, S. and M. Hoekstra (2010), "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids", *American Economic Journal: Applied Economics*, Vol. 2/1, pp. 211-228, http://dx.doi.org/10.1257/app.2.1.211. [38]

Carrell, S., M. Page and J. West (2010), "Sex and science: How professor gender perpetuates the gender gap", *Quarterly Journal of Economics*, Vol. 125/3, pp. 1101-1144, http://dx.doi.org/10.1162/qjec.2010.125.3.1101. [59]

Chapman, C. et al. (eds.) (2015), *Routledge International Handbook of Educational Effectiveness and Improvement Research: Research, Policy, and Practice*, Routledge, Abingdon; New York, NY. [41]

Chetty, R., J. Friedman and J. Rockoff (2014), "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates", *American Economic Review*, Vol. 104/9, pp. 2593-2632, http://dx.doi.org/10.1257/aer.104.9.2593. [9]

Chetty, R., J. Friedman and J. Rockoff (2014), "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood", *American Economic Review*, Vol. 104/9, pp. 2633-2679, http://dx.doi.org/10.1257/aer.104.9.2633. [10]

Collie, R., J. Shapka and N. Perry (2012), "School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy", *Journal of Educational Psychology*, Vol. 104/4, pp. 1189-1204, http://dx.doi.org/10.1037/a0029356. [49]

Darling-Hammond, L. (2017), “Teacher education around the world: What can we learn from international practice?”, *European Journal of Teacher Education*, Vol. 40/3, pp. 291-309, http://dx.doi.org/10.1080/02619768.2017.1315399. [3]

Dee, T. (2005), “A teacher like me: Does race, ethnicity, or gender matter?”, *American Economic Review*, Vol. 95/2, pp. 158-165, http://dx.doi.org/10.1257/000282805774670446. [56]

Devine, A. et al. (2012), “Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety”, *Behavioral and Brain Functions*, Vol. 8/33, pp. 1-9, http://dx.doi.org/10.1186/1744-9081-8-33. [52]

DiPrete, T. and C. Buchmann (2013), *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*, Russell Sage Foundation, New York, NY. [50]

Farb, A. and J. Matjasko (2012), “Recent advances in research on school-based extracurricular activities and adolescent development”, *Developmental Review*, Vol. 31/1, pp. 1-48, http://dx.doi.org/10.1016/j.dr.2011.10.001. [47]

Gaviria, A. and S. Raphael (2001), “School-based peer effects and juvenile behavior”, *Review of Economics and Statistics*, Vol. 83/2, pp. 257-268, http://dx.doi.org/10.1162/00346530151143798. [39]

Gershenson, S. (2016), “Linking teacher quality, student attendance, and student achievement”, *Education Finance and Policy*, Vol. 11/2, pp. 125-149, http://dx.doi.org/10.1162/EDFP_a_00180. [16]

Gneezy, U., M. Niederle and A. Rustichini (2003), “Performance in competitive environments: Gender differences”, *The Quarterly Journal of Economics*, Vol. 118/3, pp. 1049-1074, http://dx.doi.org/10.1162/00335530360698496. [53]

Grolnick, W. and M. Slowiaczek (1994), “Parents’ involvement in children’s schooling: A multidimensional conceptualization and motivational model”, *Child Development*, Vol. 65/1, pp. 237-252, http://dx.doi.org/10.2307/1131378. [64]

Hadjar, A. et al. (2014), “Gender and educational achievement”, *Educational Research*, Vol. 56, pp. 117-125, http://dx.doi.org/10.1080/00131881.2014.898908. [23]

Hallinger, P. (2018), “Bringing context out of the shadows of leadership”, *Educational Management Administration and Leadership*, Vol. 46/1, pp. 5-24, http://dx.doi.org/10.1177/1741143216670652. [42]

Hallinger, P. (2015), “The evolution of instructional leadership”, in *Assessing Instructional Leadership with the Principal Instructional Management Rating Scale*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-15533-3_1. [43]

Hanushek, E. and S. Rivkin (2010), “Generalizations about using value-added measures of teacher quality”, *American Economic Review*, Vol. 100/2, pp. 267-271, http://dx.doi.org/10.1257/aer.100.2.267. [11]

Hastie, T., R. Tibshirani and J. Friedman (2017), “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”, in *Springer Series in Statistics*, Springer, New York, https://web.stanford.edu/~hastie/ElemStatLearn//printings/ESLII_print12.pdf. [34]

Hattie, J. (2009), *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*, Routledge, London. [5]

Jackson, C. (2018), "What do test scores miss? The importance of teacher effects on non-test score outcomes", *Journal of Political Economy*, Vol. 126/5, pp. 2072-2107, http://dx.doi.org/10.1086/699018. [17]

Jackson, C., J. Rockoff and D. Staiger (2014), "Teacher effects and teacher-related policies", *Annual Review of Economics*, Vol. 6, pp. 801-825, http://dx.doi.org/10.1146/annurev-economics-080213-040845. [12]

Jennings, J. and T. DiPrete (2010), "Teacher effects on social and behavioral skills in early elementary school", *Sociology of Education*, Vol. 83/2, pp. 135-159, http://dx.doi.org/10.1177/0038040710368011. [18]

Kane, T. and D. Staiger (2008), "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation", *NBER Working Paper Series*, No. 14607, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w14607. [13]

Klassen, R. et al. (2013), "Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries", *European Journal of Psychology of Education*, Vol. 28/4, pp. 1289-1309, http://dx.doi.org/10.1007/s10212-012-0166-x. [61]

Koedel, C. (2008), "Teacher quality and dropout outcomes in a large, urban school district", *Journal of Urban Economics*, Vol. 64/3, pp. 560-572, http://dx.doi.org/10.1016/j.jue.2008.06.004. [19]

Kraft, M. (2019), "Teacher effects on complex cognitive skills and social-emotional competencies", *Journal of Human Resources*, Vol. 54/1, pp. 1-36, http://dx.doi.org/10.3368/jhr.54.1.0916.8265R3. [20]

Ladd, H. and L. Sorensen (2017), "Returns to teacher experience: Student achievement and motivation in middle school", *Education Finance and Policy*, Vol. 12/2, pp. 241-279, http://dx.doi.org/10.1162/EDFP_a_00194. [21]

LaRocque, M., I. Kleiman and S. Darling (2011), "Parental involvement: The missing link in school achievement", *Preventing School Failure: Alternative Education for Children and Youth*, Vol. 55/3, pp. 115-122, http://dx.doi.org/10.1080/10459880903472876. [63]

Le Donné, N., P. Fraser and G. Bousquet (2016), "Teaching Strategies for Instructional Quality: Insights from the TALIS-PISA Link Data", *OECD Education Working Papers*, No. 148, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jln1hlsr0lr-en. [24]

McLean, C. and E. Anderson (2009), "Brave men and timid women? A review of the gender differences in fear and anxiety", *Clinical Psychology Review*, Vol. 29/6, pp. 496-505, http://dx.doi.org/10.1016/j.cpr.2009.05.003. [54]

Mostafa, T. and J. Pál (2018), "Science teachers' satisfaction: Evidence from the PISA 2015 teacher survey", *OECD Education Working Papers*, No. 168, OECD Publishing, Paris, https://dx.doi.org/10.1787/1ecdb4e3-en. [67]

Muijs, D. et al. (2014), "State of the art: Teacher effectiveness and professional learning", *School Effectiveness and School Improvement*, Vol. 25/2, pp. 231-256, http://dx.doi.org/10.1080/09243453.2014.885451. [37]

Niederle, M. and L. Vesterlund (2011), "Gender and competition", *Annual Review of Economics*, Vol. 3, pp. 601-630, http://dx.doi.org/10.1146/annurev-economics-111809-125122. [55]

OECD (2020), *Global Teaching InSights: A Video Study of Teaching*, OECD Publishing, Paris, https://dx.doi.org/10.1787/20d6f36b-en. [66]

OECD (2020), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/ca768d40-en. [36]

OECD (2020), *PISA 2018 Technical Report*, https://www.oecd.org/pisa/data/pisa2018technicalreport/ (accessed on 16 November 2020). [33]

OECD (2020), *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/19cf08df-en. [28]

OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/5f07c754-en. [30]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b5fd1b8f-en. [31]

OECD (2019), *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/acd78851-en. [32]

OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/1d0bc92a-en. [29]

OECD (2019), *TALIS 2018 Technical Report*, OECD, Paris, http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf. [26]

OECD (2018), *Effective Teacher Policies: Insights from PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264301603-en. [4]

OECD (2017), "How do teachers teach?", *Teaching in Focus*, No. 18, OECD Publishing, Paris, http://dx.doi.org/10.1787/9a824d83-en. [69]

OECD (2015), *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264229945-en. [25]

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. [57]

OECD (2010), *Pathways to Success: How Knowledge and Skills at Age 15 Shape Future Lives in Canada*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264081925-en. [68]

Raitano, M. and F. Vona (2013), "Peer heterogeneity, school tracking and students' performances: Evidence from PISA 2006", *Applied Economics*, Vol. 45/32, pp. 4516-4532, http://dx.doi.org/10.1080/00036846.2013.791020. [45]

Rangvid, B. (2007), "School composition effects in Denmark: Quantile regression evidence from PISA 2000", *Empirical Economics*, Vol. 33, pp. 359-388, http://dx.doi.org/10.1007/s00181-007-0133-6. [46]

Rice, J. (2003), *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*, Economic Policy Institute, Washington, DC. [6]

Rivkin, S., E. Hanushek and J. Kain (2005), "Teachers, schools, and academic achievement", *Econometrica: The Journal of the Econometric Society*, Vol. 73/2, pp. 417-458, http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x. [14]

Rockoff, J. (2004), "The impact of individual teachers on student achievement: Evidence from panel data", *American Economic Review*, Vol. 94/2, pp. 247-252, http://dx.doi.org/10.1257/0002828041302244. [15]

Ruzek, E. et al. (2015), "Using value-added models to measure teacher effects on students' motivation and achievement", *The Journal of Early Adolescence*, Vol. 35/5-6, pp. 852-882, http://dx.doi.org/10.1177/0272431614525260. [22]

Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", in Hanushek, E., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1. [40]

Salisbury, J., G. Rees and S. Gorard (1999), "Accounting for the differential attainment of boys and girls at school", *School Leadership and Management*, Vol. 19/4, pp. 403-426, http://dx.doi.org/10.1080/13632439968943. [51]

Schleicher, A. (2018), *Valuing our Teachers and Raising their Status: How Communities Can Help*, International Summit on the Teaching Profession, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264292697-en. [65]

Schmidt, W., P. Zoido and L. Cogan (2014), "Schooling Matters: Opportunity to Learn in PISA 2012", *OECD Education Working Papers*, No. 95, OECD Publishing, Paris, https://dx.doi.org/10.1787/5k3v0hldmchl-en. [44]

Seidel, T. and R. Shavelson (2007), "Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results", *Review of Educational Research*, Vol. 77/4, pp. 454-499, http://dx.doi.org/10.3102/0034654307310317. [7]

Tschannen-Moran, M. and A. Hoy (2001), "Teacher efficacy: Capturing an elusive construct", *Teaching and Teacher Education*, Vol. 17/7, pp. 783-805, https://doi.org/10.1016/S0742-051X(01)00036-1. [62]

Van De Werfhorst, H. and J. Mijs (2010), "Achievement inequality and the institutional structure of educational systems: A comparative perspective", *Annual Review of Sociology*, Vol. 36, pp. 407-28, http://dx.doi.org/10.1146/annurev.soc.012809.102538. [1]

Viac, C. and P. Fraser (2020), "Teachers' well-being: A framework for data collection and analysis", *OECD Education Working Papers*, No. 213, OECD Publishing, Paris, https://dx.doi.org/10.1787/c36fc9d3-en. [48]

Wayne, A. and P. Youngs (2003), "Teacher characteristics and student achievement gains: A review", *Review of Educational Research*, Vol. 73, pp. 89-122, https://www.jstor.org/stable/3516044. [8]

Notes

[1] Teacher effect refers to the systematic variation in outcome across students assigned to the same teacher given the teacher's ability to increase students' initial knowledge and skills, which accounts for significant variation in student achievement.

[2] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[3] In each round of PISA, one subject, among reading, mathematics and science, is tested in detail. The main subject in 2018 was reading.

[4] In Malta, there are only 17 out of the 44 schools that are not single-gender schools (i.e. all students surveyed in the school are same-gender students) and where the within-school differences in performance between girls and boys can be computed. Thus, it is not included in the school-level analysis presented in Chapter 4.

[5] Since Viet Nam does not have data on PISA test scores, it is not included in the analyses presented in Chapters 2 and 4.

[6] The depth of survey questions not only matters for the validity of the instrument in terms of its precision, it also matters for its cross-cultural comparability. Longer questions with many sub-items seem indeed to be a better option to enhance cross-cultural comparability of measurement (Avvisati, Le Donné and Paccagnella, 2019[70]).

[7] Analysis conducted on previous TALIS-PISA link data (2012-13) showed that there is also a high degree of alignment between teachers and students' perceptions of teachers' classroom practices (OECD, 2017[69]).

[8] This is consistent with past findings established based on PISA 2015 data, with science as the major domain (Mostafa and Pál, 2018[67]). In eight of the 19 countries and economies that participated in the PISA 2015 teacher questionnaire, students in schools with more-satisfied science teachers performed better in science than students in schools with less-satisfied science teachers. Yet, after accounting for students' and schools' socio-economic profile, the relation was significant in fewer countries and economies.

[9] More information regarding how local communities can help schools and their teachers can be found in the report by Schleicher, and in particular in Chapter 2 (Schleicher, 2018[65]).

[10] Malta and Viet Nam were not included in this analysis. See Endnotes 3 and 4 above.

[11] This is a national option already offered in PISA, which is quite popular among participating countries and economies.

[12] As shown by an OECD report, the Canadian example has demonstrated the value of linking PISA to a longitudinal follow-up, which can improve our understanding of the social and economic impact of competencies acquired by the school-going population (OECD, 2010[68]).

2 What do teachers and schools do that matters most for student achievement?

We know that teachers and schools matter. However, there is less certainty about the specific characteristics and actions of teachers and school leaders that matter for student achievement. This chapter explores teacher and school factors that are significantly related to student achievement in the three subject domains covered by PISA: reading, mathematics and science. In order to best harness the richness of the TALIS-PISA link data, the analysis is centred around a machine learning technique. While the chapter focuses mainly on the characteristics and practices of teachers and schools that matter for student performance in all three subjects, it also attempts to identify cross-country patterns, differential teacher and school effects and the mediating effects of classmates' characteristics.

Highlights

- TALIS-PISA link[1] data are more likely to provide insights for the Czech Republic and Turkey, where differences in school average performances represent about half of the total variance in student achievements, as opposed to countries, including Australia, Denmark and Malta, where 25% or less of the total variation in student outcomes lie between schools.
- A machine learning technique applied to TALIS-PISA link data retains six potential key predictors of student achievement in reading, mathematics and science: teachers' classroom practices, teachers' well-being and job satisfaction, teachers' use of working time, classmates' characteristics, school culture and school leadership. Variance decomposition analysis suggests that each of these factors explain at least 20% of the variation in student performances between schools.
- On average across the countries and economies participating in the TALIS-PISA link, students tend to perform better in all three subjects covered by PISA (i.e. reading, mathematics and science) the more class time the average school teacher spends on actual teaching and learning. Moreover, results suggest that the use of class time is more closely linked to student performance in mathematics than in reading and science.
- As teachers are more satisfied with their work environment, students tend to perform better in school on average across TALIS-PISA link countries and economies. Moreover, results suggest that teachers' satisfaction with their work environment is more closely related to student performance in science than in reading and mathematics. Students who attend schools where teachers report workload as an important source of stress tend to perform better in school on average across the countries and economies participating in the TALIS-PISA link. This may signal teachers' commitment and dedication to their work as well as highly competitive school environments (attended by higher-performing students) that can lead to workload being an important source of stress for teachers.
- As the amount of working hours teachers spend on marking and correcting increases on average within the school students tend to perform better academically on average across the countries and economies participating in the TALIS-PISA link. This may point to the importance of assessment of student work in providing feedback to students about their learning progress. Yet, it may also signal highly competitive school environments (attended by higher-performing students) that can lead to more frequent feedback to students in the form of tests and exams.
- TALIS-PISA link data suggest the presence of peer effects. Indeed, as the average concentration of students from socio-economically disadvantaged homes in the classrooms increases, students tend to perform worse academically in several countries and economies participating in the TALIS-PISA link. This finding holds while accounting for students' own socio-economic background. In addition, in most countries and economies participating in the TALIS-PISA link, as the average concentration of academically gifted students in the classrooms increases, the better students tend to perform. This may not only signal the presence of peer effects, but also the presence of academic segregation.
- On average across TALIS-PISA link countries and economies, students who attend schools where stakeholders (i.e. parents and local community) are involved in school-related activities tend to perform better in school. However, this association does not remain significant for most participating countries and economies once classmates' characteristics are taken into account.
- The findings highlighted above cannot be interpreted as causal but only as correlational given that TALIS and PISA measure student, teacher, principal and school characteristics in many countries at a single point of time. For example, while students may perform better if a greater share of class time is spent on actual learning, the causality can also go the other way around. Indeed, more disruptive classrooms are more likely to have lower-achieving students, which, in turn, leads to more time spent on other tasks such as keeping order or administrative tasks.

Introduction

Identifying the factors that help younger generations to succeed later in life has long been a main interest for education policy. In fact, student achievement is driven by many factors that are beyond the reach of the school, including students' abilities and attitudes, family background and support as well as peer group effects (OECD, 2005[1]). Although teachers and school leaders also have an impact on student learning outcomes by shaping the quality of instruction (Barber and Mourshed, 2009[2]; Darling-Hammond, 2017[3]; OECD, 2018[4]), the evidence is less clear about the specific characteristics of teachers and school leaders that matter for student achievement. Therefore, investigating what teachers and schools do that matter for student performance can provide insights to raise the quality of education and ensure that every student succeeds in school and later in life.

Past educational research has shown that teacher quality is the most important school-related predictor of student achievement (Hattie, 2009[5]; Rice, 2003[6]; Seidel and Shavelson, 2007[7]; Wayne and Youngs, 2003[8]). It is well established that teacher effect (or teacher's value-added), which refers to the systematic variation in outcome across students assigned to the same teacher given the teacher's ability to increase students' initial knowledge and skills, accounts for significant variation in student achievement (Chetty, Friedman and Rockoff, 2014[9]; Chetty, Friedman and Rockoff, 2014[10]; Hanushek and Rivkin, 2010[11]; Jackson, Rockoff and Staiger, 2014[12]; Kane and Staiger, 2008[13]; Rivkin, Hanushek and Kain, 2005[14]; Rockoff, 2004[15]). The magnitude of teacher effect estimates is relatively large compared to the effects of other school factors (Jackson, Rockoff and Staiger, 2014[12]).[2]

There is evidence that teacher's value-added goes beyond test scores. For example, Chetty, Friedman and Rockoff (2014[10]) found that students exposed to high-quality teachers (i.e. those with high value-added based on their impacts on students' test scores) in primary school are more likely to attend college, earn higher salaries, live in neighbourhoods with a higher share of college graduates, and have higher savings rates. Girls who had good teachers in primary school are also less likely to have teenage pregnancies. Good teachers also have an impact on social-emotional competencies, such as students' beliefs in their mathematics skills, behaviour in class, perseverance, growth mindset, happiness and truancy (Blazar and Kraft, 2017[16]; Jackson, 2018[17]; Kraft, 2019[18]).

Thus, we do know that teachers matter. However, evidence is less conclusive about the specific characteristics and actions of teachers that matter for student achievement. Observable teacher characteristics, such as level of education and type of certification, tend to explain little of the variation in teacher effectiveness (Gordon, Kane and Staiger, 2008[19]; Hanushek and Rivkin, 2010[11]; Jackson, Rockoff and Staiger, 2014[12]; Kane, Rockoff and Staiger, 2008[20]; Rivkin, Hanushek and Kain, 2005[14]). Another body of research focusing on attributes of new teaching candidates show that no single teacher characteristic predicts student achievement (Dobbie, 2011[21]; Rockoff et al., 2011[22]). However, the combination of certain attributes – such as general intelligence, personality traits, beliefs regarding self-efficacy (Rockoff et al., 2011[22]); and academic achievement, leadership experience, perseverance, critical thinking, organisational ability, motivational ability and respect for others (Dobbie, 2011[21]) – explain a substantial part of the variance in teachers' value-added.

The complexity and context-based nature of a teacher's work provide some explanation for the fact that research still strives to pinpoint what makes an effective teacher. One strand in the literature on teacher effectiveness focuses on the relationship between teachers' actual classroom practices and student outcomes (Hattie, 2009[5]; Muijs et al., 2014[23]). One of the main findings from this strand of research is that teachers' classroom practices explain a large share of the variance in student outcomes (Muijs et al., 2014[23]). What happens in the classroom is important for student performance – the learning environment in the classroom, the quantity, pacing and quality of instruction, and interaction with students (Muijs et al., 2014[23]; Teddlie and Reynolds, 2000[24]). In addition, there is a distinct line of research that moves away from indirect measures of teaching, typically in the form of questionnaires, and focuses on direct measures like classroom observation through video. One such example is the TALIS Video Study, which is an international large-scale, video-based study of teaching and learning featuring a longitudinal design. It shows that the quality of instructional practices applied in the classroom has the largest impact on student learning compared to classroom management and socio-emotional practices, before accounting for students' abilities and background (OECD, 2020[25]).

In turn, teachers' classroom practices are embedded in the context and functioning of the school. School-level factors that determine the context and functioning of the school have mainly indirect effects on student performance by developing and evaluating the school policy on teaching and the policy on creating a learning environment at the school (Creemers and Kyriakides, 2010[26]). School effectiveness research, which tries to establish the school-level determinants of students' learning outcomes, indicates relatively small effects from schools' resource-input and organisational measures (Scheerens, 2001[27]). However, schools' instructional conditions are found to have average to large effects on school effectiveness (Scheerens, 2001[27]). For instance, Dobbie and Fryer (2013[28]) finds that the school-level factors that matter for school effectiveness include frequent teacher feedback, data-driven instruction, high-dosage tutoring, increased instructional time and a relentless focus on academic achievement, as opposed to input measures such as class size, per-pupil expenditure and the fraction of teachers with an advanced degree.

The teacher and school factors that matter for student achievement also vary depending on the different student outcomes. Research findings on teacher effectiveness (Muijs et al., 2014[23]; Rice, 2003[6]; Seidel and Shavelson, 2007[7]; Wayne and Youngs, 2003[8]) and school effectiveness (Hattie, 2009[5]; Reynolds et al., 2014[29]) point to the notion of multidimensionality of teacher and school effectiveness. Indeed, there is empirical evidence for differential teacher and school effects across subjects (Muijs et al., 2014[23]; Reynolds et al., 2014[29]; Rockoff, 2004[15]; Seidel and Shavelson, 2007[7]). Thus, analysing the teacher- and school-level factors that matter for student achievement by subject domain may provide more nuanced insights that are specific to student outcomes in a given subject.

Chapter 2 of this report attempts to identify specific characteristics and practices of teachers and school leaders that matter the most for student achievement. It explores these in the three subject domains covered by PISA:[3] reading, mathematics and science. It draws on the rich TALIS-PISA 2018 link dataset of 15-year-old students and their teachers and schools, and presents results for eight countries and economies – Australia, Ciudad Autónoma de Buenos Aires (hereafter CABA [Argentina]), Colombia, the Czech Republic, Denmark, Georgia, Malta and Turkey.[4] The first section discusses how teacher and school dimensions can be mapped to student achievement and how they interact with each other. The chapter then presents the analytical approach taken to best harness the richness of the TALIS-PISA link data. The third section presents the teacher and school factors that matter for student achievement across the three subject domains covered by PISA, given the characteristics of the TALIS-PISA link data and in light of the modelling approach taken. This section not only explores the characteristics and practices of teachers and schools that are significantly related to student performance in all three subjects, but also attempts to identify cross-country patterns, differential teacher and school effects and the mediating effects of classmates' characteristics.

Conceptual mapping of teacher and school factors to student achievement

The theoretical framework of this chapter discusses the approach taken to conceptualise the relationship between teacher and school factors, and student cognitive outcomes. First, it presents a framework with two axes of effect and level, based on which the relationship between the dimensions of teacher and school characteristics and practices can be linked to student achievement. Second, it introduces the teacher and school dimensions that are included in the analysis by focusing on previous research findings about the potential effect of a given dimension on students' cognitive outcomes. Last, it provides a rationale for the subject domain focus of the analysis and for the inclusion of certain student characteristics as controls.

The conceptual mapping of Teaching and Learning International Survey (TALIS) dimensions to student achievement is framed along two axes: effect and level (Figure 2.1). The first axis (i.e. effect) is based on whether a dimension can have a direct effect on student achievement. For example, the student composition of the classroom, or certain elements of school culture, such as parental support and involvement, can have a direct effect on student outcomes irrespective of teachers' classroom practices. That said, most of the teacher and school factors collected by TALIS influence student achievement indirectly through their effect on the quality of instruction. The second axis (i.e. level) refers to whether the dimension mainly operates at the school[5] or teacher level. Dimensions that reveal more about the characteristics of a school and that are mostly influenced by school- or system-level policies are categorised as school-level dimensions. However, teachers' practices that are specific to certain teachers in a given school and are mainly at the discretion of individual teachers belong to the teacher level. Yet, it is important to note that the TALIS-PISA link data only allows for the link between teacher-level dimensions and student achievement at the school level. Hence, dimensions that reflect on the characteristics and practices of teachers measured at the school level are referred to as teacher dimensions (or factors).

This analysis focuses on the association between teacher and school characteristics and practices, and student outcomes. Nevertheless, it is important to note that relationships between teacher and school dimensions are often characterised by reciprocity and inter-connectedness. For example, professional development influences classroom practices, and in turn, those practices have an effect on the type of professional development provided to teachers (Ainley and Carstens, 2018[30]). Moreover, these feedback loops are present across different levels. For instance, not only do school culture and school leadership influence classroom practices but they themselves are shaped by the practices teachers use in the classroom. Indeed, the various levels and components of the educational system are inter-related and work in interaction (Reynolds et al., 2014[29]). Therefore, dynamic models of school effectiveness aim to capture these reciprocal associations between variables and their effect on student outcomes in an integrated and comprehensive fashion (Creemers and Kyriakides, 2015[31]; Creemers and Kyriakides, 2010[26]; Creemers and Kyriakides, 2007[32]; Kyriakides, Christoforou and Charalambous, 2013[33]).

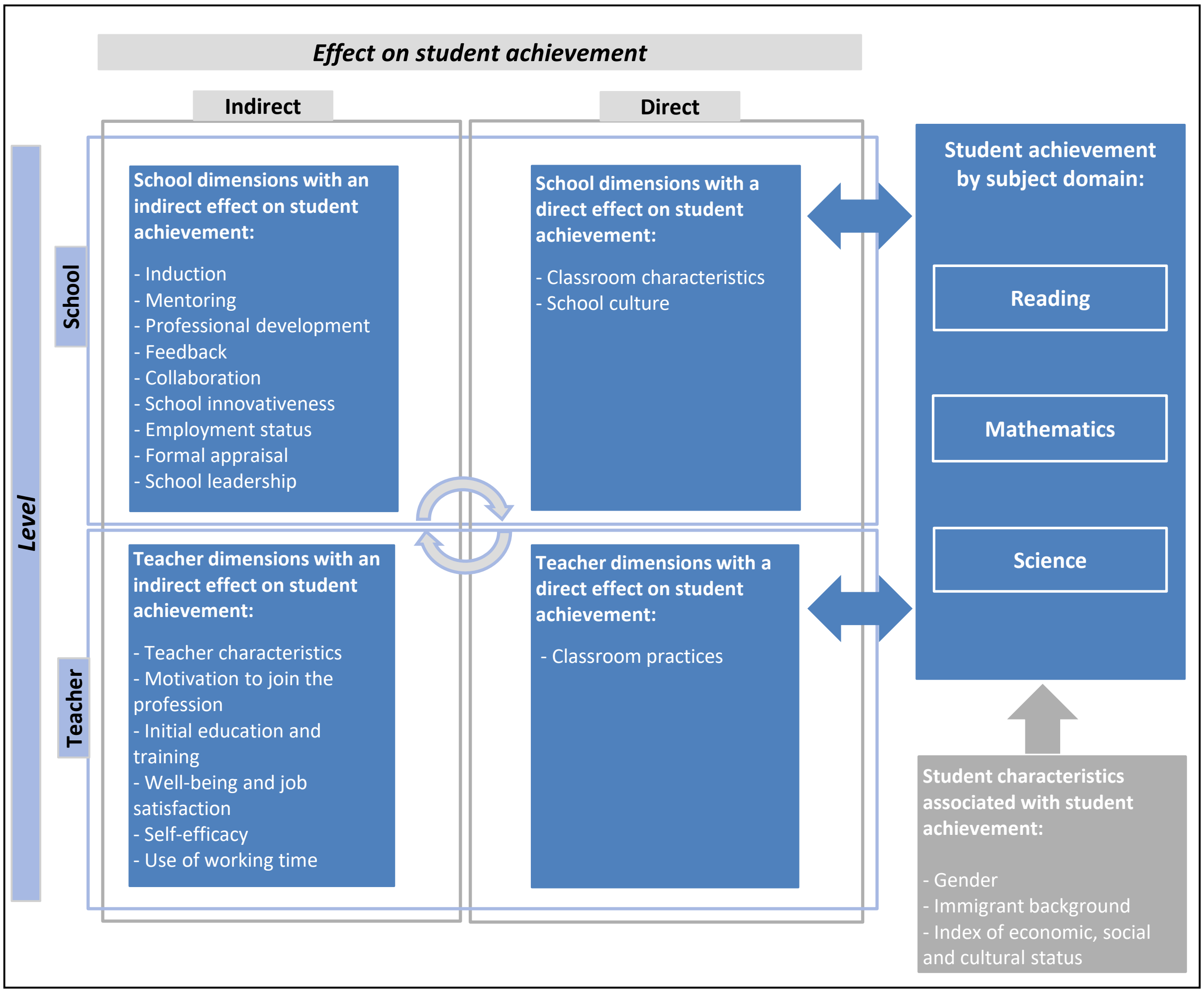
Given the often reciprocal nature of relationships, certain factors can be both an input and an output of schooling (Creemers and Kyriakides, 2015[31]). Indeed, the reciprocity also holds for the relationship between various teacher and school factors, and student achievement. For instance, student performance can have an impact on the choice of teaching strategies applied in the classroom, but it can also influence other factors such as school culture (e.g. teacher-student relations), the type of professional development provided to teachers or teacher well-being, job satisfaction and self-efficacy.

Teacher dimensions with a direct effect on student achievement

Classroom practices

The importance of teachers' classroom practices is key to any study of teaching and learning because what teachers do in the classroom is the strongest direct school-based influence on student learning outcomes (Hattie, 2009[5]). Most other school factors influence student learning mainly because they influence teachers' practices and, thereby, have a transmitted influence on student learning (OECD, 2019[34]). Teachers' classroom practices encompass various dimensions, including classroom management, teacher support, clarity of instruction, cognitive activation and feedback to students (Ainley and Carstens, 2018[30]). According to past research, the core factors that may be significantly related to student outcomes include, among others, the quantity and pacing of instruction (i.e. use of class time, classroom disciplinary climate, academic pressure) and the modalities of instruction and interaction (i.e. giving information, asking questions, providing feedback) (Muijs et al., 2014[23]). In this report, the dimension of teachers' classroom practices includes indicators on use of class time, autonomy over planning and teaching, teachers' perceived disciplinary climate,[6] use of practices related to clarity of instruction and cognitive activation as well as use of different types of assessment practices.[7]

Figure 2.1. The conceptual mapping of TALIS dimensions to student achievement



Research looking at the relationship between different teaching strategies and students' academic achievement demonstrate that certain strategies are more closely associated with student performance than others. A former OECD study exploring the relationship between student performance on the PISA mathematics test and exposure to specific teaching approaches found that higher-achieving students reported being more often exposed to cognitive-activation instruction[8] (Echazarra et al., 2016[35]). Teaching practices that focus on cognitive activation consist of exposing students to problems for which there is no immediately obvious solution or that can be solved in several different ways, or by using alternative routes and procedures to solve a problem. These instructional practices may also require students to apply what they have learned to new contexts (Echazarra et al., 2016[35]). Similarly, the previous 2013 TALIS-PISA link report (Le Donné, Fraser and Bousquet, 2016[36]), which focused on the relationship between mathematics teachers' teaching strategies and student learning outcomes in mathematics as well as students' attitudes toward learning, showed that, overall, a frequent use of the cognitive activation strategy and, to a lesser extent, the active learning strategies, were associated with higher mathematics performances. In contrast with cognitive activation, teacher-directed instruction consists of practices based on lecturing, memorisation and repetition, where the teacher is the main actor responsible for transmitting knowledge to receptive students (Le Donné, Fraser and Bousquet, 2016[36]).

Teacher dimensions with an indirect effect on student achievement

Teacher characteristics, motivation to join the profession and initial education and training

The evidence on the relationship between observable teacher characteristics, such as level of education, type of certification and experience, and student achievement is mixed. For example, Rivkin, Hanushek and Kain (2005[14]) found no evidence that a master's degree improves teacher effectiveness. Regarding teacher certification, recent OECD findings suggest that students in schools with a larger share of certified teachers perform better in reading, as measured by PISA (OECD, 2020[37]). However, other research findings show that the type of teaching certification has negligible impact on student test scores (Kane, Rockoff and Staiger, 2008[20]). The evidence on the relationship between experience and teacher effectiveness is also mixed. While there is research showing that the benefits of additional years of experience tend to fade away after the initial years in the profession (Rivkin, Hanushek and Kain, 2005[14]), there is also evidence that experience matters in teacher effectiveness later in the career (Papay and Kraft, 2015[38]).

Teachers' characteristics, background and initial education and training may have an effect on student outcomes due to their influence on the quality of teachers' instruction. For instance, it is plausible to assume that novice teachers build up their skills and become more effective in their teaching as they become more experienced, which eventually translates into better student outcomes. Moreover, the knowledge acquired during initial education and training has an effect on the teaching strategies adopted by teachers and the quality of their instruction (Blömeke, Gustafsson and Shavelson, 2015[39]), which are, in turn, significantly related to student achievement (Baumert et al., 2010[40]; Hill, Rowan and Ball, 2005[41]; Kersting et al., 2012[42]).

Teachers' motivation to join the profession can influence student achievement through indirect channels. The reasons motivating someone to join the teaching profession strongly correlate with the extent of job satisfaction (Ainley and Carstens, 2018[30]). Moreover, teachers' desire to make teaching their first choice as a profession also relates to job satisfaction and reported self-efficacy (OECD, 2019[34]). Teachers whose first career choice was teaching are more likely to be satisfied with their job in almost all countries and economies participating in TALIS, and they also tend to report higher self-efficacy in around two-thirds of countries and economies participating in TALIS (OECD, 2019[34]).

Job satisfaction, well-being and self-efficacy

Teachers' job satisfaction, well-being and beliefs are inter-related and have an indirect effect on student outcomes. Job satisfaction, which is the sense of fulfilment and gratification from working, has a positive impact on teachers, school culture and, ultimately, on students (Ainley and Carstens, 2018[30]). Based on TALIS results, teachers' job satisfaction (both with the current work environment and with the profession) is positively associated with teachers' self-efficacy (OECD, 2020[43]; OECD, 2014[44]). Moreover, job satisfaction also plays a key role in teachers' attitudes, efforts and confidence in their daily work with students (Caprara et al., 2003[45]; Klassen et al., 2013[46]; Tschannen-Moran and Hoy, 2001[47]). Thus, positive job satisfaction may also have an indirect effect on student achievement through enhanced commitment leading to better performance of teachers.

Well-being and stress, whether classroom- or workload-based, is integral to teachers' job satisfaction (Ainley and Carstens, 2018[30]). Stressful working environments and challenging working conditions can affect the practices of teachers and principals, their motivation for their work and even student achievement (Viac and Fraser, 2020[48]). Indeed, research has associated high levels of stress with lower self-efficacy for teaching, lower job satisfaction, and lower commitment (Collie, Shapka and Perry, 2012[49]). Conversely, teachers with high levels of well-being are likely to report higher levels of self-efficacy and job satisfaction as well as stronger motivation at work (Viac and Fraser, 2020[48]).

There is consensus among educational researchers, policy makers and practitioners that teacher self-efficacy is an essential teacher characteristic. It is strongly associated with teachers' pedagogical practices and the quality of teachers' instruction (Holzberger, Philipp and Kunter, 2013[50]), which are, in turn, associated with students' academic achievement. (Ainley and Carstens, 2018[30]). Moreover, teachers with high self-efficacy show higher job satisfaction and commitment, and are less likely to be affected by burnout (Avanzi et al., 2013[51]; Chesnut and Burley, 2015[52]; Klusmann et al., 2008[53]; Skaalvik and Skaalvik, 2010[54]).

Use of working time

The way teachers balance their time among competing tasks is important for the quality of the teaching and for student learning (OECD, 2019[34]). Apart from actual teaching, teachers also need to allocate their working time to other tasks, including, among others, lesson planning and preparation, marking and correcting student work, general administrative work, team work and dialogue with colleagues as well as communication and co-operation with parents. The way the different tasks of teachers are related to student learning varies. Allocating more time to administrative work allows less time for core activities such as lesson planning and preparation, which directly affect student learning in the classroom. The time teachers spend on preparing their lessons is highly beneficial to the quality of their instruction (Boeskens and Nusche, 2021[55]; Hargreaves, 1992[56]; Paniagua and Istance, 2018[57]).

Teachers' use of working time may not only influence student outcomes through their effect on the quality of teaching, but also through teachers' stress and well-being. Indeed, research shows that time pressures and workload are among the main factors affecting teachers' stress and well-being (Bakker et al., 2007[58]; Collie, Shapka and Perry, 2012[49]; Hakanen, Bakker and Schaufeli, 2006[59]; Klassen and Chiu, 2010[60]; Viac and Fraser, 2020[48]). Based on TALIS data, teachers' stress is just slightly related to teaching intensity (OECD, 2020[43]). However, the share of teachers reporting that they experience a great deal of stress in their work increases more sharply in planning, marking and, particularly, administrative tasks (OECD, 2020[43]).

School dimensions with a direct effect on student achievement

Classroom characteristics

The classroom's student composition in terms of cultural background, language spoken at home, socio-economic background, ability level and learning needs, can strongly influence student outcomes. Research on school-level peer effects show that the performances of both low- and high-achieving students are negatively affected by the presence of classmates with learning difficulties. Yet, high-performing students tend to be less affected than their low-achieving peers by the composition of their classes (Burke and Sass, 2013[61]; Lavy, Silva and Weinhardt, 2012[62]; Sacerdote, 2011[63]). There are various channels – both direct and indirect – through which peer effects can influence student outcomes. Students not only learn directly from their classmates with higher innate ability, but competition with high-achieving peers often makes students more motivated and work harder (Sacerdote, 2011[63]). Peer effects may also have a positive effect on future education expectations and career aspirations of students who lack sufficient knowledge and hold lower ambitions compared to their socio-economically more advantaged peers (OECD, 2019[64]). On the negative side, peer effects can influence student achievement indirectly through reduced teaching time due to disruptive behaviour or the detrimental effect of teaching practices being adapted to the needs of low performers (OECD, 2019[64]; Sacerdote, 2011[63]). The previous TALIS-PISA link report showed that the type of teaching strategies applied matters more for students' mathematics performance in socio-economically advantaged schools than in disadvantaged ones (Le Donné, Fraser and Bousquet, 2016[36]).

It sounds plausible that class size, another attribute of the classroom, has an inverse relationship with students' academic achievement since teachers can devote more time and attention to each student. Indeed, TALIS data show that smaller classes tend correlate with more actual teaching and learning time (OECD, 2019[34]). PISA 2018 results show that, on average, education systems with smaller language-of-instruction classes tend to feature higher PISA reading scores (OECD, 2020[37]). However, class size is not a predictor for other quality indicators of teaching processes captured by TALIS, such as the use of cognitive activation practices and teachers' reported self-efficacy in teaching (OECD, 2019[34]). Indeed, the empirical evidence on the effects of class size reduction on student achievement is somewhat mixed with certain research findings suggesting very little or no effect (Hanushek, 2006[65]; Hanushek, 2003[66]; Hanushek, 1999[67]; Hoxby, 2000[68]) while others find evidence of a positive impact (Angrist and Lavy, 1999[69]; Krueger, 2003[70]). Nevertheless, overall, the positive effects of smaller classes, especially at primary level, seem to be established by the literature (Bouguen, Grenet and Gurgand, 2017[71]). And research is also conclusive that, due to the high costs of such policy intervention, there can be a trade-off between reducing class size and improving teaching quality (Jepsen and Rivkin, 2009[72]; Rivkin, Hanushek and Kain, 2005[14]).

School culture

School culture is a multi-faceted concept that includes participation of staff, teachers and students in school decisions, collaborative school culture, student-teacher relations, academic pressure and involvement of parents and the community. This has both direct and indirect effects on students. The right school culture can enhance student learning directly and foster the conditions for effective teacher instruction (Ainley and Carstens, 2018[30]). For instance, parental support of student achievement, which is one of the factors captured by stakeholders' involvement, can have a direct effect on student outcomes irrespective of teachers' classroom practices. Indeed, past research findings show that a positive school culture has a powerful influence on many of the elements that affect both students and teachers. School culture relates not only to student learning and well-being (Battistich et al., 1997[73]; Bryk and Schneider, 2002[74]; Cohen et al., 2009[75]; Engel, Rutkowski and Rutkowski, 2009[76]; Hoy, Tarter and Hoy, 2006[77]; Martin et al., 2013[78]; Nilsen and Gustafsson, 2014[79]), but also to teacher effectiveness, confidence, and commitment to teaching (Hoy and Woolfolk, 1993[80]; Thapa et al., 2013[81]; Weiss, 1999[82]). Regarding the specific elements of school culture that may have a particularly close relationship with student achievement, Muijs et al. (2014[23]) highlight student-teacher relations and teacher expectations of students.

School dimensions with an indirect effect on student achievement

Teacher induction, mentoring, professional development and feedback

Support and early professional development is crucial for the development of novice teachers, who not only lack experience but often face more challenging working conditions than their more experienced peers (OECD, 2019[34]). Empirical evidence shows that support and assistance for beginning teachers in the form of induction and mentoring have a positive influence on outcomes such as commitment and retention of teachers, classroom teaching practices and student achievement (Ingersoll and Strong, 2011[83]). Induction and mentoring are also identified at the system level as being common to high-performing and equitable education systems (OECD, 2018[4]). Teacher induction can have an effect on student outcomes through teaching quality and job satisfaction. TALIS data show that teachers who took part in some kind of induction activity, formal or informal, also tend to report higher self-efficacy and job satisfaction on average across OECD countries and economies (OECD, 2019[34]). Rockoff (2008[84]) provides evidence of the beneficial effects of high-intensity mentoring on teaching quality and, ultimately, student achievement.

Supporting teachers to improve their skills is equally important at later stages of a teacher's career. Since teachers are assumed to be lifelong learners with different professional needs throughout their careers it is important that their instructional practices and ability to implement innovation in teaching and learning are developed through continuous professional development (Ainley and Carstens, 2018[30]; OECD, 2019[34]). Indeed, effective continuous professional development programmes can improve teachers' skills (Borko, 2004[85]; Garet et al., 2016[86]) and classroom practices (Fischer et al., 2018[87]), thereby influencing student outcomes. For example, there is empirical evidence that a high-intensity and multifaceted in-service teacher training programme, which is designed to improve teaching skills (rather than provide course content) and includes counselling and feedback sessions for teachers, is associated with students' test score improvements (Angrist and Lavy, 2001[88]). Indeed, research concludes that professional development programmes are effective when they are intensive, sustained, collaborative and focused on teachers' practice (Opfer, 2016[89]). Moreover, as a recent OECD study shows, one of the common components of high-performing education systems is the opportunity for in-service teachers to participate in professional development workshops organised by the school (OECD, 2018[4]).

Feedback to teachers in the form of evaluation and appraisal can affect teaching quality. It is considered a key feature of effective professional development (Ingvarson, Meiers and Beavis, 2005[90]) and continuous learning (Jensen and Reichl, 2011[91]). By recognising teachers' strengths and addressing weaknesses in their pedagogical practices, feedback improves teachers' effectiveness (OECD, 2014[44]; OECD, 2013[92]), and has the largest impact on student performance of any school intervention (Hattie, 2009[5]). TALIS 2018 results show a significant positive association between teachers' job satisfaction and receiving impactful feedback (OECD, 2020[43]). Moreover, peer feedback from other teachers is a particularly important and unique form of collaboration between educators as it involves close contact and interaction between colleagues, driven by the purpose of learning from colleagues' expertise and suggestions (OECD, 2020[43]).

Collaboration

Research points to the value of two or more teachers interacting or working together to accomplish a specific goal. Teacher collaboration enables teacher learning, stemming from the exchange of ideas and interactions (Goddard, Goddard and Tschannen-Moran, 2007[93]). It helps teachers learn from each others' practices and experiences, which can help improve their own practices (Reeves, Pun and Chung, 2017[94]). Based on TALIS data, teachers who report deeper forms of collaboration (also referred to as professional collaboration) with their peers – such as team teaching, providing feedback based on classroom observations, engaging in joint activities across different classes and participating in collaborative professional learning – tend to report higher levels of job satisfaction, self-efficacy and the use of cognitive activation practices in the classroom (OECD, 2020[43]). Thus, teacher collaboration may have an indirect effect on student outcomes through its influence on teachers' instructional practices, job satisfaction and self-efficacy.

Except for a few countries and economies that participated in TALIS, most of the variation in teachers' responses regarding deeper forms of collaborative activities is spread across teachers (OECD, 2020[43]). This suggests that a teacher who collaborates within a school does not collaborate with all teachers of the school but only a few, while other colleagues from the same school do not collaborate at all (OECD, 2020[43]). Nevertheless, teacher collaboration is considered a school factor since school leadership has a major role in shaping the degree of collaboration as well as the culture of collaboration in the school. Indeed, research shows that leadership actions of school leaders are strong predictors of collaborative actions between teachers (Leithwood, Leonard and Sharratt, 1998[95]; Marks and Printy, 2003[96]; O'Donnell and White, 2005[97]). Moreover, analysis based on TALIS data suggests that teachers who report that their school provides staff with opportunities to participate in school decisions also tend to engage in deeper forms of collaborative activities more frequently (OECD, 2020[43]).

School innovativeness

Teachers' openness to change and their take-up of innovative teaching practices is important in supporting students' acquisition of cross-curricular skills. Today's generation of students needs broader and more complex skills – creativity, innovation, problem solving, critical thinking and digital literacy – to succeed in our complex and rapidly changing world (Ainley and Carstens, 2018[30]). But teachers must first embrace these skills in their teaching before they can develop them in their students (Ainley and Carstens, 2018[30]). Research evidence shows that teacher openness and extraversion – qualities that lend themselves to innovativeness – significantly influence teacher performance in the classroom (Ainley and Carstens, 2018[30]). TALIS aims to capture teachers' innovativeness through teachers' perceptions of their colleagues' openness towards innovation and how conducive the school context is toward innovation. In practice, innovativeness within TALIS refers to innovation at the school level,[9] and is therefore categorised as a school dimension.

Employment status and formal appraisal

Working conditions significantly affect the quality of the teacher's job and, subsequently, the teacher's motivation, engagement and well-being as well as the learning environment (Bascia and Rottmann, 2011[98]; Gomendio, 2017[99]; Viac and Fraser, 2020[48]). Job insecurity in the form of fixed-term employment, for example, can lead to insecurity and unpredictability, which may cause strain and prevent certain teachers from functioning optimally in their work environment (de Cuyper, de Witte and Van Emmerik, 2011[100]). Flexible time arrangements such as part-time work (both voluntary and involuntary), can help teachers achieve work-life balance and personal well-being. However, they can also have a negative impact on career progression and professional practices, including collaboration with other colleagues (Ainley and Carstens, 2018[30]; OECD, 2020[43]). Teachers' employment status can be categorised as a teacher dimension as different teachers in the same school may have different employment arrangements. Yet, for the theoretical framework of this chapter, employment status is defined as a school dimension since it also reflects on system-level characteristics of an education system.

Previous TALIS findings provide some insights into the different ways employment status influence student outcomes through its effect on teacher practices. Based on TALIS 2013 results, fixed-term employment is associated with participating less in formal induction programmes and professional development activities, and receiving less mentoring (OECD, 2014[44]). Similarly, results from TALIS 2018 show that part-time teachers are less likely to participate in professional development (OECD, 2019[34]) or professional collaboration (OECD, 2020[43]). And short-term work and part-time work tend to be negatively associated with teachers' self-efficacy (OECD, 2020[43]).

Teacher appraisal, which refers to the formal evaluation of teachers, is an important element of high-performing schools (Ainley and Carstens, 2018[30]; OECD, 2020[43]). In fact, a common characteristic of high-performing education systems is the presence of teacher-appraisal mechanisms, either legislated or deeply rooted in school practice, with a strong focus on teachers' continuous improvement (OECD, 2018[4]).Teacher appraisal can be a tool for ensuring that required standards are met or recommended practices followed. It also allows teachers to reflect on their teaching practice – their strengths and weaknesses – and identify areas for improvement (OECD, 2020[43]). Indeed, well-designed appraisal systems can support effective teaching practices (Ainley and Carstens, 2018[30]). They can also improve recognition of teachers' efforts and competencies, leading to more satisfied and motivated teachers (Isoré, 2009[101]; OECD, 2013[92]). Thus, teacher appraisal may have an indirect effect on student achievement through teachers' practices and beliefs.

School leadership

School leadership is one of the most important school factors influencing students' achievement (Chapman et al., 2015[102]; Hallinger, 2018[103]). The relationship between leadership and student outcomes is considered indirect as effective leadership practices aim to create supportive learning environments in which teachers are able to develop their practices and engage effectively with students' learning (Hallinger, 2011[104]; Muijs, 2011[105]; Reynolds and Muijs, 2016[106]). Instructional leadership is of particular interest in the context of the relationship between school leadership and student outcomes since it refers to principals' efforts to improve teachers' instructional quality (Ainley and Carstens, 2018[30]; OECD, 2020[43]). Instructional leadership refers to principals' actions to manage curriculum, attend to teachers' professional development needs and create a culture of collaboration (Hallinger, 2015[107]; Hallinger, 2011[104]; Hallinger and Heck, 2010[108]), and there is empirical evidence of its effect on student learning outcomes (Goddard et al., 2015[109]; Hallinger, 2015[107]).

Differential effects on student achievement by subject domain

Former research found differences in the size and nature of teacher and school effects across subjects. While teaching experience has been found to be more beneficial for test scores in reading than in mathematics (Rockoff, 2004[15]), teacher and school effects in mathematics and science tend to be larger than in reading (Seidel and Shavelson, 2007[7]; Reynolds et al., 2014[29]). One explanation for these findings is that, unlike reading, mathematics and science are mainly learned at school while their exposure in the family is more limited (Reynolds et al., 2014[29]). Although effective schools in one subject area tend to be more effective in other areas, there is also evidence for the contrary showing that certain schools can be effective in one subject domain but ineffective in another (Reynolds et al., 2014[29]). Therefore, the research literature warrants that the identification of the teacher and school factors that matter for student achievement be conducted by subject domain. In particular, as TALIS collects information on the subjects taught by teachers, it is possible to analyse student performances in a given subject by looking specifically at what the corresponding subject teachers do.[10]

Student characteristics associated with student achievement

Besides the teacher and school factors that are the main point of interest in this chapter, it is also important to take into account that there are student-level characteristics influencing student achievement. Ignoring them may lead to wrong conclusions about the relationship between teacher and school characteristics and practices, and student outcomes if there are student-level factors that are significantly related to student achievement but which are omitted from the analysis. Indeed, based on evidence from PISA, there are certain student characteristics, including student's socio-economic background, gender and immigrant background, which are reliable predictors of student outcomes. There is firm research evidence showing that the most stable predictor of a student's future success at school is his or her family background (OECD, 2019[64]). In addition, PISA results suggest that students' gender can also predict academic performance. Girls tend to outperform boys in reading and, to a lesser extent, boys outperform girls in mathematics on average across all participating countries and economies (OECD, 2019[64]).[11] The general pattern of the relationship between students' immigrant status and their achievement on the PISA test show that non-immigrant students outperform their first and second-generation immigrant peers in most countries (OECD, 2019[64]). However, this association is closely related to the socio-economic background of immigrant students, and thus may not hold for countries that apply selective immigration policies favouring highly skilled immigrants.

How to make the most of TALIS-PISA link data in exploring the factors that matter for student achievement?

Linking TALIS data, which provides information regarding the background, beliefs and practices of teachers and principals, and PISA data, which delivers insights into the background characteristics and cognitive and non-cognitive skills of 15-year-old students, offers an internationally comparable dataset combining information on key stakeholders. The breadth and depth of the TALIS-PISA link data provides an opportunity to identify teacher and school factors that have a significant association with student achievement.

Yet, the TALIS-PISA link data also present important limitations. The link between teachers and students is established at the school and not at the classroom level. In other words, the data do not allow matching a teacher with her or his students; rather, the data only allow matching a sample of teachers teaching 15-year-old students in a school with a sample of 15-year-old students of that same school. Information on teachers is therefore averaged at the school level and then analysed together with students' outcomes. Given that teachers of the same school differ significantly in terms of their characteristics and practices, linking data by averaging teachers' variables at the school level triggers a considerable loss of information.[12] Moreover, the cross-sectional design of the survey prevents measurement of any causal effects of teachers. Neither can short-term effects of teachers and schools on students' outcomes be distinguished from long-term ones.

As the focus of this chapter is to examine the teacher and school factors that matter for student achievement, TALIS and PISA data are linked by merging individual student data collected by PISA with TALIS principal data and TALIS teacher data averaged at the school level. Due to the survey design of the TALIS-PISA link data discussed above, teacher dimensions measure the average teacher's characteristics and practices at the school. Nevertheless, analysing the relationship between teacher and school factors, and student performance at the student level allows student characteristics at the level of students to be accounted for.

In order to make the most of the TALIS-PISA link data, this chapter takes a mixed data- and theory-driven approach. It is data-driven insofar as the identification of the teacher and school characteristics and practices that matter for student outcomes is centred around a supervised statistical learning method. And it is theory-driven because theory and previous research findings further inform the selection of variables used for the statistical method as well as interpret and validate the findings.

The main statistical technique used to investigate the relationship between teacher and school factors, and student achievement is the least absolute shrinkage and selection operator (also known as lasso). It is a machine learning technique within the family of supervised statistical learning methods (Box 2.1). Lasso is an attractive tool in analysing the relationship between the many variables collected through the TALIS questionnaires and the student outcomes measured by the PISA test. It selects the variables that are highly correlated with the outcome variable even when the number of potential variables is high relative to the number of observations (Box 2.1).

Box 2.1. The least absolute shrinkage and selection operator (lasso)

The least absolute shrinkage and selection operator (also known as lasso) is a machine learning technique within the family of supervised statistical learning methods. It has several attributes that enable it to identify the teacher and school factors that matter for student outcomes out of the many potential variables collected through the TALIS questionnaires. These are:

- Lasso is designed to select variables that are important and should be included in the model.
- The outcome variable guides the model selection process (i.e. supervised statistical learning method).
- Lasso can handle high-dimensional models where the number of variables is high relative to the number of observations.

Lasso estimates coefficients in a model. It selects variables that correlate well with the outcome in one dataset (training sample) and then tests whether the selected variables predict the outcome well in another dataset (validation sample). In the TALIS-PISA link data analysis, the training and validation samples are sub-samples of the TALIS-PISA link sample. Lasso proceeds with model selection by estimating model coefficients in such a way that some of the coefficient estimates are exactly zero, and hence, excluded from the model, while others are not (Hastie, Tibshirani and Friedman, 2017[110]; Hastie, Tibshirani and Wainwright, 2015[111]; Tibshirani, 1996[112]).

While lasso has several attributes that makes it an attractive tool for model selection, there are certain assumptions and limitations that need to be taken into account when applying this statistical technique:

- Lasso is most useful when only a few out of many potential variables affect the outcome (Hastie, Tibshirani and Friedman, 2017[110]; Hastie, Tibshirani and Wainwright, 2015[111]; Tibshirani, 1996[112]). The assumption that the number of coefficients that are non-zero (i.e. correlated with the outcome variable) in the true model is small relative to the sample size is known as a sparsity assumption. The approximate sparsity assumption requires that the number of non-zero coefficients in the model that best approximates the true model be small relative to the sample size.
- In the context of model selection, lasso may not always be able to distinguish an irrelevant predictor that is highly correlated with the predictors in the true model from the true predictors.
- Applying lasso for model selection means finding a model that fits the data, not finding a model that allows for interpreting estimated coefficients as effects. Thus, when used for model selection, lasso selects variables and estimates coefficients but it does not provide the standard errors required for performing statistical inference.

Note: For additional information on lasso, see Annex B.
Sources: Hastie, T., R. Tibshirani and J. Friedman (2017[110]), "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", in *Springer Series in Statistics*, https://web.stanford.edu/~hastie/ElemStatLearn//printings/ESLII_print12.pdf; Hastie, T., R. Tibshirani and M. Wainwright (2015[111]), "Statistical learning with sparsity: The lasso and generalizations", *Monographs on Statistics and Applied Probability*, https://web.stanford.edu/~hastie/StatLearnSparsity_files/SLS.pdf; Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society: Series B (Methodological)*, http://dx.doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Nevertheless, lasso also has certain limitations (Box 2.1) that warrant a balanced approach to identifying the teacher and school dimensions that might be key predictors of student achievement. Therefore, the results from lasso regressions are complemented with a theory-driven approach grounded in the relevant research literature and a variance decomposition analysis.

Identifying the dimensions that explain most of the differences in school average performances (and ultimately, variance in student achievement) may provide insights about specific teacher and school factors that may be significantly related to student performance. Yet, the variance decomposition analysis complements the findings from the lasso regressions as it can reveal the relative importance of each teacher and school dimension in explaining the average differences in student performances across schools. Moreover, unlike the results of lasso regressions (Box 2.1), the variance decomposition analysis is not influenced by the presence of highly correlated variables.

Once the teacher and school dimensions that might be key predictors of student achievement are identified based on lasso regression results, the variance decomposition analysis, as well as the relevant research literature, then the country-level regression analyses (featuring teacher and school dimensions separately) are applied. Thus, students' individual PISA scores are regressed on indicators of each teacher and school dimension (taken separately) that lasso regression results flag as important predictors. These dimensions explain a substantial part of between-school variance in student performance (i.e. 20% or above on average across subjects). In the context of this chapter, multiple linear regressions are estimated on one dimension at a time and provide insights into how the value of student performance in a given subject changes when any one of the independent variables within a dimension varies while all other independent variables included in the model are held constant. In comparison to lasso regressions, multiple linear regressions provide the confidence intervals of the coefficient estimates, which, in turn, allow for drawing inferences about the overall population. Moreover, they lead to more accurate coefficient estimates through the introduction of final and balanced repeated replicate weights and the use of plausible values of student performance. In contrast to lasso regressions, which are based on the overall population of 15-year-old students, teachers and principals surveyed within the TALIS-PISA link, multiple linear regressions are applied at the country level. This allows cross-country patterns to be established. Nested multiple linear regressions provide additional insights into the effects of potential confounding factors such as classmates' characteristics, and the relationship between teacher- or school-level factors and student achievement. Furthermore, the country-level multiple linear regressions do not focus solely on subject teachers but also looks at the relationships of interest by considering all teachers within the school. This provides the opportunity to identify differential teacher and school effects on student outcomes across subject domains.

What are the characteristics and practices of teachers and schools that matter for student achievement across subject domains?

Past research has shown that teacher and school factors do matter for student achievement. However, the evidence is mixed regarding the specific characteristics and practices of teachers and principals that are significantly related to students' cognitive outcomes across the three subject domains covered by PISA – reading, mathematics and science. The TALIS-PISA link data may provide further insights into this field of research given that it is an internationally comparable dataset combining information on students, teachers and principals.

This section aims to investigate the teacher and school factors that matter for student achievement across the three subject domains covered by PISA. First, it presents the dimensions that are selected by lasso regression analysis as being significantly related to student performance in all three subjects. Apart from presenting the teacher and school dimensions that are consistently highlighted by lasso, the first section also discusses the indicators selected for at least two out of the three subject domains within each highlighted dimension. Then, the section examines which dimensions explain a substantial share of the variance in student achievement at the school level. Lastly, it looks at the relationships between student achievement and the dimensions that are highlighted by the mixed data- and theory-driven approach more in detail, focusing on cross-country patterns, differential teacher and school effects in relation to the three subjects covered and the mediating effects of classmates' characteristics.

Teacher and school factors selected by lasso, a supervised statistical learning method

Following a data-driven approach, one can let lasso regressions select which of the almost 150 potential predictors of the 18 teacher and school dimensions actually matter for student performance in reading, mathematics and science. Figure 2.2 provides an overview of the teacher and school factors that are significantly related to student performance by focusing on the average subject domain teacher within schools.[13] Lasso regression results suggest that, considering the overall population of 15-year-old students, subject domain teachers and principals surveyed within the TALIS-PISA link, a third of the teacher and school dimensions under consideration matter for student performance consistently in all three subject domains covered by PISA.

The six dimensions highlighted consistently by lasso regression analysis as having a significant association with student achievement in all three subjects include both teacher and school dimensions as well as factors with direct and indirect effects on student achievement (Figure 2.2). These dimensions are teachers' classroom practices, teachers' well-being and job satisfaction, teachers' use of working time, classmates' characteristics, school culture and school leadership.

Teachers' classroom practices

The relationship between classroom practices and student achievement is well established in past research (Hattie, 2009[5]; Le Donné, Fraser and Bousquet, 2016[36]; Muijs et al., 2014[23]). Among the various aspects of teachers' classroom practices, such as teachers' autonomy over planning and teaching, teaching strategies, classroom management and assessment practices, lasso regression results highlight the use of class time and classroom disciplinary climate[14] as factors that are significantly related to student performance in at least two out of the three subject domains (Figure 2.2). Indeed, the link between the quantity and pacing of instruction, and student achievement is consistently confirmed by past research (Muijs et al., 2014[23]). Research has shown that students' opportunity to learn has significant effects on student achievement (Schmidt, Zoido and Cogan, 2014[113]). Clearly, students' opportunities to learn is closely linked to the amount of time allocated to academic instruction (Muijs et al., 2014[23]). However, the amount of time students are actively engaged in learning during a lesson is, in turn, highly related to the classroom's disciplinary environment and the teacher's classroom management practices and skills (Muijs et al., 2014[23]). As previous TALIS findings show, experienced teachers spend more time on actual teaching and learning partly because they teach less challenging classrooms (OECD, 2019[34]). Hence, the positive association highlighted by lasso between the share of class time spent on actual teaching and learning – as opposed to administrative and classroom management tasks – and student achievement may also point to a reverse causal relationship. Indeed, more disruptive classrooms are more likely to have lower achieving students, which, in turn, leads to more time spent on other tasks such as keeping order or administrative tasks.

Teachers' well-being and job satisfaction

Lasso regression results are in line with previous research findings, which indicate that teachers' satisfaction with their job can have an indirect positive effect on student achievement through teachers' beliefs, attitudes and practices as well as school culture (Ainley and Carstens, 2018[30]). In particular, teachers' job satisfaction with work environment is found to have a significant positive relationship with student performance (Figure 2.2). Yet, this finding may also point to reverse causality. It can be assumed that teachers may be particularly satisfied with their school assignment when they work in schools attended by high-achieving students. These teachers might be particularly committed to helping their academically gifted students progress further, and, in turn, the students may respond positively to the teachers' increased motivation.

Figure 2.2. Teacher and school factors that matter for student achievement, based on lasso

Dimensions and variables selected by lasso regressions based on responses of the overall population of 15-year-old students, subject domain teachers and principals surveyed within the TALIS-PISA link, by subject

Dimensions	Variables selected	Reading	Maths	Science
		(sign of standardised coefficients: +/-)		
Teacher dimensions with a direct effect on student achievement				
Classroom practices	Use of class time (%): Keeping order in the classroom	-	-	
	Use of class time (%): Actual teaching and learning	+	+	+
	Teachers' perceived disciplinary climate *(higher values indicate higher need for classroom discipline)*		-	-
Teacher dimensions with an indirect effect on student achievement				
Teacher characteristics	Age		+	
Motivation to join the profession	*No variables selected*			
Initial education and training	Highest level of formal education completed: ISCED 8	+		+
	Content: Teaching in a mixed ability setting	-	-	
	Content: Monitoring students' development and learning	-	-	
	Sense of preparedness: Content of some or all subject(s) I teach		+	
	Sense of preparedness: Teaching in a mixed ability setting			-
Well-being and job satisfaction	Workload stress *(higher values indicate workload being considered a more important source of stress)*	+	+	
	Job satisfaction with work environment	+	+	+
	Job satisfaction with profession		-	-
	Teachers' views of the way different stakeholders value the profession	-	-	
Self-efficacy	*No variables selected*			
Use of working time	Working hours by tasks: Marking/correcting of student work	+	+	+
	Working hours by tasks: General administrative work			-
	Working hours by tasks: Professional development activities	-		
School dimensions with a direct effect on student achievement				
Classroom characteristics (*classmates' characteristics and class size*)	Classroom composition (%): Low academic achievers	-	-	-
	Classroom composition (%): Students with special needs	-	-	
	Classroom composition (%): Students with behavioural problems		-	
	Classroom composition (%): Students from socio-economically disadvantaged homes	-	-	-
	Classroom composition (%): Academically gifted students		+	+
	Class size (no. of students)	+		+
School culture	Teachers' actions towards achieving academic excellence		+	
	Stakeholder (i.e., parents and local community) involvement in school-related activities	+	+	+
	Teacher-student relations	+	+	
School dimensions with an indirect effect on student achievement				
Induction	Took part in formal induction activities during first employment			+
Mentoring	*No variables selected*			
Professional development	Type: Online courses/seminars	+		
	Type: Observation visits to other schools	-		-
	Type: Participation in a network of teachers formed specifically for the professional development of teachers			-
	Type: Reading professional literature		+	
	Content: Knowledge and understanding of my subject field(s)	+		
	Content: School management and administration			-
	Content: Approaches to individualised learning	-		-
	Content: Teaching students with special needs	-	-	
	Content: Teacher-parent/guardian co-operation			-
Feedback	Feedback received by source: External individuals or bodies	+	-	
	Feedback received by method: School-based and classroom-based results (e.g. performance results, project results, test scores)	+		
Collaboration	*No variables selected*			
School innovativeness	*No variables selected*			
Employment status	Permanent employment		+	
	Part-time (50-70% of full time hours)		-	
	Part-time (less than 50% of full time hours)	-		
Formal appraisal	*No variables selected*			
School leadership	Principals' leadership activities: Collaborated with teachers to solve classroom discipline problems	-		-
	Principals' leadership activities: Took actions to support co-operation among teachers to develop new teaching practices	-	-	-
	Principals' leadership activities: Collaborated with principals from other schools on challenging work tasks		-	

Notes: Country fixed effects and student characteristics, such as gender, immigrant background and index of economic, social and cultural status are always included among the variables selected by lasso.
Teacher variables are averaged at the school level for subject domain teachers. Thus, the analyses by subject are based on samples restricted to schools with at least one subject domain teacher.
Dimensions and variables selected consistently across subject domains are highlighted in blue. Dimensions that are not selected based on any of the subject domains are highlighted in light grey. Since lasso is applied as a model selection technique, it does not provide the standard errors required for performing statistical inference. The interpretation of the estimated standardised coefficients is conditional on the selected model and cannot be interpreted as causal. Moreover, in the presence of correlated explanatory variables, the signs of the coefficient estimates can swing based on which other independent variables are in the model.
For additional information on the full list of potential variables included in the lasso regressions, as well as more information on lasso in general, see Annex B.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 2.1.

Lasso regression results also highlight teachers' workload-induced stress[15] as an important factor in student performance (Figure 2.2). Interestingly, the relationship between workload being an important source of stress and student achievement is positive. Workload-induced stress may signal teachers' commitment and dedication to their work. While teachers' passion and dedication to their work can be a source of satisfaction, it can also lead to burnout by engaging with the work with too much intensity (Carbonneau et al., 2008[114]). But, this finding may also point to a reverse causal relationship as highly competitive school environments (attended by higher-performing students) can lead to workload being an important source of stress for teachers.

Based on lasso regression analysis, there is a significant negative relationship between average teachers' satisfaction with the profession and student achievement (Figure 2.2). TALIS asks teachers how they generally feel about their profession – whether the advantages of being a teacher outweigh the disadvantages; if they would still choose to work as a teacher if they could decide again; if they regret that they decided to become teachers; and if they wonder whether it would have been better to choose another profession. The negative relationship with student achievement is observed specifically in the case of student performance in mathematics and science. This could be explained by the possibility that mathematics and science teachers who teach in schools where student achievement is high may be more apt to consider alternative opportunities on the labour market that have better career opportunities and remuneration packages. Research evidence shows that science teachers are more likely to leave the teaching profession than their peers (Allen and Sims, 2017[115]), and that increasing science teachers' salaries can have a positive effect on retention (Clotfelter et al., 2008[116]; Feng and Sass, 2018[117]). Hence, the negative association highlighted by lasso regression results likely reveals how mathematics and science teachers regard their profession in light of other potential career opportunities. Nevertheless, this finding may also result from the limitations of the lasso regressions. Indeed, an irrelevant predictor may just as well be flagged by lasso as an important one missed (Box 2.1).

Past research found evidence for the positive relationship between teachers' social status and student achievement (Dolton et al., 2018[118]). Nevertheless, lasso regression results point towards a significant negative association between average teachers' views of the way different stakeholders value the profession[16] and student achievement (Figure 2.2). A possible explanation for this finding is that teachers in low-performing schools are actually higher up in the social ladder than most of the local adult population surrounding these schools. Higher local social status likely goes hand in hand with teachers' greater influence on local issues in comparison to other workers. This could trigger a better perception of the influence of teachers as a whole. However, similar to the lasso results related to teachers' satisfaction with the profession, this finding may also result from the limitations of the lasso regressions.

Teachers' use of working time

As part of teachers' use of working time, time spent on marking and correcting student work is selected by lasso regressions as an important predictor of student achievement (Figure 2.2). Assessment of student work, which tends to be one of the most time-consuming activities in teachers' work after teaching and lesson preparation (OECD, 2019[34]), is an essential part of a teacher's job to nurture student academic growth. Research shows that effective teaching includes formative assessment in the form of constructive and immediate feedback and that this type of feedback has positive implications for teaching and learning (Muijs and Reynolds, 2001[119]). Nevertheless, summative assessment of students' work, in the form of tests and exams, also provides feedback to students about their learning progress (Ainley and Carstens, 2018[30]). In addition, teachers' time spent on summative assessment can be indicative of their engagement in formative assessment. Thus, teachers' working hours spent on marking and correcting student work can be linked to student performance through teachers' feedback on students' learning progress. Nevertheless, it can be assumed that the relationship between hours spent marking and correcting, and student performance is non-linear. Providing feedback to students in the form of tests and exams only benefits learning outcomes if it does not lead to students' exhaustion and burnout. In addition, this finding may also point to a reverse causal relationship as highly competitive school environments (attended by higher- performing students) can provide more frequent feedback to students in the form of tests and exams.

Classmates' characteristics and class size

As one would expect, lasso regression analysis points to a significant relationship between the context in which teaching and learning takes place in schools, and student outcomes (Figure 2.2). The relationship between student achievement and the average concentration of low academic achievers, students with special needs and students from socio-economically disadvantaged backgrounds within classrooms is negative, while the association between student performance and the concentration of academically gifted students in classes is positive.[17] Indeed, the existence of peer effects in relation to student outcomes is well established in the research literature (Ammermueller and Pischke, 2009[120]; Avvisati et al., 2014[121]; Burke and Sass, 2013[61]; Chetty, Hendren and Katz, 2016[122]; Duflo, Dupas and Kremer, 2011[123]; Lavy, Silva and Weinhardt, 2012[62]; Sacerdote, 2011[63]). Past analyses of PISA data also show that students, regardless of their own socio-economic background, tend to be advantaged scholastically if they attend a school whose students are from more advantaged socio-economic backgrounds (OECD, 2013[124]; OECD, 2019[64]). Students can potentially influence the motivations, the behaviour and ultimately the academic performance of their peers directly through interaction (Avvisati et al., 2014[121]; Sacerdote, 2011[63]). Moreover, students can also have an indirect effect on their classmates through the overall level of teacher effort and teachers' choice of the level at which to target instruction (Duflo, Dupas and Kremer, 2011[123]; Sacerdote, 2011[63]).

Based on lasso regression results, there is a significant positive relationship between the average class size at a school and student achievement (Figure 2.2). However, the relationship between class size and student achievement is hard to disentangle. Research shows that smaller classes tend to have a positive effect on student achievement (Bouguen, Grenet and Gurgand, 2017[71]). However, academically gifted students tend to attend larger classes. Typically, schools sort the least able students into smaller classes, which results in well-performing students attending larger classes (Bouguen, Grenet and Gurgand, 2017[71]). As a result, the true effect of class size on pupil performance cannot be determined without accounting for selection bias due to the way students are usually sorted into different classes.

School culture

Besides classroom characteristics, lasso regression analysis also highlights the importance of school culture in relation to student achievement (Figure 2.2). Past research findings show that school culture can have both direct and indirect influences on students and teachers, and as a result it is closely related to student achievement (Ainley and Carstens, 2018[30]). In particular, findings from lasso regressions point to the involvement of parents and the community[18] as a specific factor in a positive relationship with student performance. Indeed, research evidence is conclusive on the positive association between parental and community involvement, and student achievement (Wang and Degol, 2016[125]; Wilder, 2014[126]). Former research provides some insights regarding the channels through which parental involvement may be linked to student performance. The relationship tends to be the strongest if parental involvement refers to higher expectations for academic achievement, as opposed to involvement being defined as homework assistance (Wilder, 2014[126]). In addition, it can be also assumed that the positive association between stakeholder involvement in school-related activities and student achievement partly captures the overall positive influence of parental involvement in an adolescent's life, including psychological and emotional support that are not necessarily observed by school leaders. Moreover, there is evidence for spillover effects in parents' involvement in education as the impact of more involved parents on their children is amplified at the class level by peer group interaction (Avvisati et al., 2014[121]). Nevertheless, similar to most of the relationships presented above, this finding may also signal reverse causality. As students perform better academically, parents may see more worth in supporting students' achievement and being more involved in school-related activities.

In addition to stakeholder involvement, lasso regression results indicate that teacher-student relations[19] are another important element of school culture with respect to student performance (Figure 2.2). This finding is in line with previous research that identified teacher-student interaction as an important component in creating an effective learning environment (Muijs et al., 2014[23]). Together with the finding that shows the importance of the share of class time spent on actual teaching and learning, the potential link between student-teacher relations and student achievement highlights the importance of the classroom-level context when it comes to student performance.

School leadership

Past research findings identify school leadership as one of the most important school factors influencing students' achievement (Chapman et al., 2015[102]; Hallinger, 2018[103]). Lasso regression analysis highlights, in particular, principals' actions to support co-operation among teachers in developing new teaching practices as an important predictor of student performance, but the association is negative (Figure 2.2). Similarly, according to lasso regression analysis, there is a negative relationship between student performance and the frequency with which principals collaborate with teachers to solve classroom disciplinary problems, and with principals from other schools to solve challenging work tasks. These findings may signal a reverse causal relationship. It can be assumed that in school environments characterised by co-operation among teachers as well as a lack of disciplinary issues and challenging work tasks, which also happen to be attended by higher-performing students, school leaders do not feel the need to engage in such activities on a regular basis.

Teacher and school dimensions not selected by lasso

More than half of the teacher and school dimensions introduced in the regressions are not consistently selected across subjects as important predictors by lasso. These dimensions are assumed to have indirect links with student achievement (Figure 2.1). Teachers' initial education and training, and teachers' participation in professional development are flagged by lasso as important predictors of student performance in each subject (Figure 2.2). However, there is no common variable within these two dimensions of teacher training deemed relevant for all of reading, mathematics and science. Moreover, most of the variables selected by lasso regressions within these two dimensions reveal a negative relationship with student achievement. For instance, lasso regression results show a negative association between the share of teachers within schools for whom teaching in a mixed ability setting was included in their formal education or who felt prepared for this type of teaching, and student achievement (Figure 2.2). What this result potentially reveals is that teachers who had formal pre-service training in teaching in challenging classrooms end up teaching such classes later on in their career. In turn, these classes tend to be characterised by lower student performance on average. Similarly, most of the associations between teachers' participation in professional development and student performance that are highlighted by lasso regression results show an inverse relationship. These negative associations potentially identify areas in which teachers are facing challenges that eventually create the link between participating in certain professional development activities and teaching low-achieving students.

Nevertheless, there are some elements within initial education and training and professional development that are selected by the lasso regressions for a specific subject, including teachers' sense of preparedness for the content of the subject taught, participation in professional development in the form of online courses or seminars and in the form of reading professional literature, and participation in professional development focusing on knowledge and understanding of subject field (Figure 2.2). It could be assumed that these factors are more closely related to student performance in a given subject as compared to other subjects. Yet, the country-level regression results by each dimension taken separately do not confirm the differential effect of teachers' pre- and in-service training across subject domains, since the relationships between the above-mentioned elements within initial education and training and student performance, as well as professional development and student performance, are not significant at the country level. (Tables 2.21, 2.22, 2.23, 2.24, 2.71, 2.72, 2.73, 2.74, 2.75, 2.76, 2.77, 2.78, 2.79, 2.80, 2.81 and 2.82). Given the limitations of the lasso regressions and/or the TALIS-PISA link data, an important factor may be missed for a given subject. In the case of specific teacher and school factors that are not consistently selected across subject domains it is not possible to draw a conclusion with respect to their relationship with student performance.

Teacher characteristics, induction, feedback and employment status are only selected for one or two subjects of the three, which are reading, mathematics and science (Figure 2.2). Since all these dimensions are assumed to have an indirect effect on student achievement (Figure 2.1), lasso regression results may simply point towards this indirect link with student outcomes. Indeed, teacher characteristics, induction, feedback and employment status may influence student performance through their effects on classroom practices and school culture, which have been incorporated in the lasso model. It could be also assumed that some of these teacher and school factors may be more closely linked to a specific subject. Yet, looking at each dimension taken separately at the country-level (Tables 2.17, 2.18, 2.19, 2.20, 2.67, 2.68, 2.69, 2.70, 2.83, 2.84, 2.85, 2.86, 2.87, 2.88, 2.89 and 2.90), as well as the average between-school differences in student performances explained by these dimensions (Figure 2.4), there is little evidence for the presence of differential teacher effects.

There are teacher and school dimensions that are not selected by lasso regressions for any of the subjects included in the analysis. These dimensions include motivation to join the profession, self-efficacy, mentoring, collaboration, school innovativeness and formal appraisal (Figure 2.2). Nevertheless, these teacher and school factors may still matter for student achievement. First, these dimensions may have an indirect effect on student achievement (Figure 2.1). For instance, teachers who report higher self-efficacy tend to be more satisfied with their job and more committed to it (Avanzi et al., 2013[51]; Chesnut and Burley, 2015[52]; Mostafa and Pál, 2018[127]; OECD, 2019[34]), while students taught by teachers who are more satisfied with their jobs and are more committed also tend to perform better in school (Ainley and Carstens, 2018[30]; Carbonneau et al., 2008[114]). As another example, engaging in professional collaboration more often may influence student achievement through the beneficial effects of the exchange of ideas on teachers' instructional practices as well as on the risk of teacher attrition (OECD, 2020[43]). Second, the lack of significant relationship may be an artefact resulting from the characteristics and limitations of the TALIS-PISA link data and/or the modelling approach taken. The aggregation of teacher data at the school level means that all the relationships analysed within this chapter relate to a school's overall context and needs to be interpreted accordingly. Therefore, it is plausible to assume that the analyses within this chapter may miss some factors that do contribute to student achievement but are not captured because no direct link can be drawn between an average teacher in a school and a sample of students.

Differences in student achievement explained by teacher and school factors

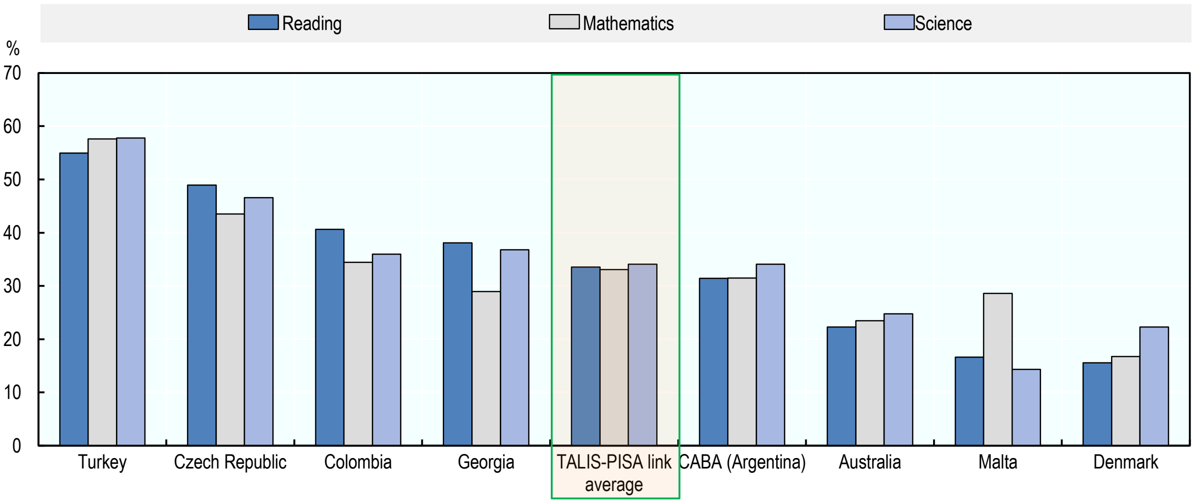
This chapter aims at identifying the teacher and school dimensions that matter the most for student performance. Analysing results from lasso regressions is one way to do this (see previous section). Another complementary approach consists of retaining those dimensions that explain the highest shares of variance in student performance by relying on a variance decomposition analysis. Although this approach cannot highlight the specific factors within a dimension that may matter the most for student achievement, it complements the findings of the lasso regressions by revealing the relative importance of each teacher and school dimension in explaining the average differences in student performance across schools. Another advantage of the variance decomposition analysis is that, unlike the results of lasso regressions (Box 2.1), the presence of highly correlated variables does not affect the results.

However, the variance decomposition analysis is not without limitations either. Most importantly, the shares of between-school variance explained by each dimension may be artificially driven by the number of variables included in a given dimension. Indeed, the dimensions that have the lowest explanatory power, such as collaboration, school innovativeness, mentoring, motivation to join the profession, employment status, self-efficacy, induction and teacher characteristics, include four or fewer variables, while the dimensions that explain larger shares of the differences in school average performance include between seven and 14 variables. Thus, caution is warranted when interpreting these results.

Yet, with the aforementioned advantages and limitations in mind, one can examine the share of variation in student achievement that is possible to explain with the help of TALIS-PISA link data as well as the explanatory power of each teacher and school dimension. TALIS-PISA link data suggest that differences in school average performance across the eight TALIS-PISA link countries and economies with available data represent about one-third of the total variance in student performance, irrespective of the subject domain (Figure 2.3). This means that on average across TALIS-PISA link countries and economies around a third of the total variation in student performance can actually be captured by the analyses included in this chapter. Obviously, as the factors of interest included in the analyses are introduced at the school level, one can only explain the share of variance in students' performance at the school level. For the same reason, as the link between teachers and students is established at the school and not at the classroom level, the within-school variance in student performance explained by each teacher dimension cannot be examined based on the TALIS-PISA link data.

Figure 2.3. Schools differ in student achievement

Percentage of total variance in PISA scores explained at the school level, by subject



Notes: The analyses are based on samples restricted to schools with at least one subject domain teacher. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Countries and economies are ranked in descending order of the percentage of total variance in PISA reading score, explained at the school level.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.2, 2.3 and 2.4.

StatLink https://doi.org/10.1787/888934223707

Nevertheless, the variation in the share of total variance in student performance that can be explained at the school level varies across countries and economies. While in the Czech Republic and Turkey differences in school average performance represent about half of the total variance in student achievements, at the other end of the spectrum, the same shares are below 25% in most subjects in Australia, Denmark and Malta (Figure 2.3). This means that the TALIS-PISA link data is more likely to provide insights for countries such as the Czech Republic and Turkey, where the school-level variations in student performance are considerably higher as opposed to countries, including Australia, Denmark and Malta, where TALIS-PISA link data can only capture 25% or less of the total variation in student outcomes.

The share of between-school variance explained by each dimension included in the analyses is fairly similar across reading, mathematics and science. In line with research findings that highlight the importance of the classroom level in relation to teaching and learning (Hattie, 2009[5]; Muijs et al., 2014[23]), classroom characteristics explain the largest share of variance in student performance that exists between schools (44-45%) (Figure 2.4). At the other end of the spectrum, school innovativeness and teacher collaboration have the lowest explanatory power. Depending on the subject, these dimensions account for 5% or less of the variation in student achievement that exists between schools (Figure 2.4).

Except for a few exceptions, the dimensions that explain the largest shares of between-school variance turn out to be the ones that are also identified by the lasso regression analysis. These nine dimensions, which represent both teacher and school dimensions as well as factors with direct and indirect effects on student achievement include: classroom characteristics (44-45%), use of working time (26-31%), school leadership (23-25%), classroom practices (23-28%), well-being and job satisfaction (17-27%) and school culture (20-22%). The dimensions that are not consistently highlighted by the lasso regressions but do explain a substantial share of between-school variance in student achievement (i.e. 20% or above on average across subjects) include professional development – both in terms of type and content (19-34%) – initial education and training – both in terms of content and teachers' sense of preparedness for a given element of pre-service training (20-27%) – and formal appraisal (19-24%) (Figure 2.4). On average across countries and economies participating in the TALIS-PISA link, these dimensions explain between 17-45% of the differences in school average performances across subjects (Figure 2.4).

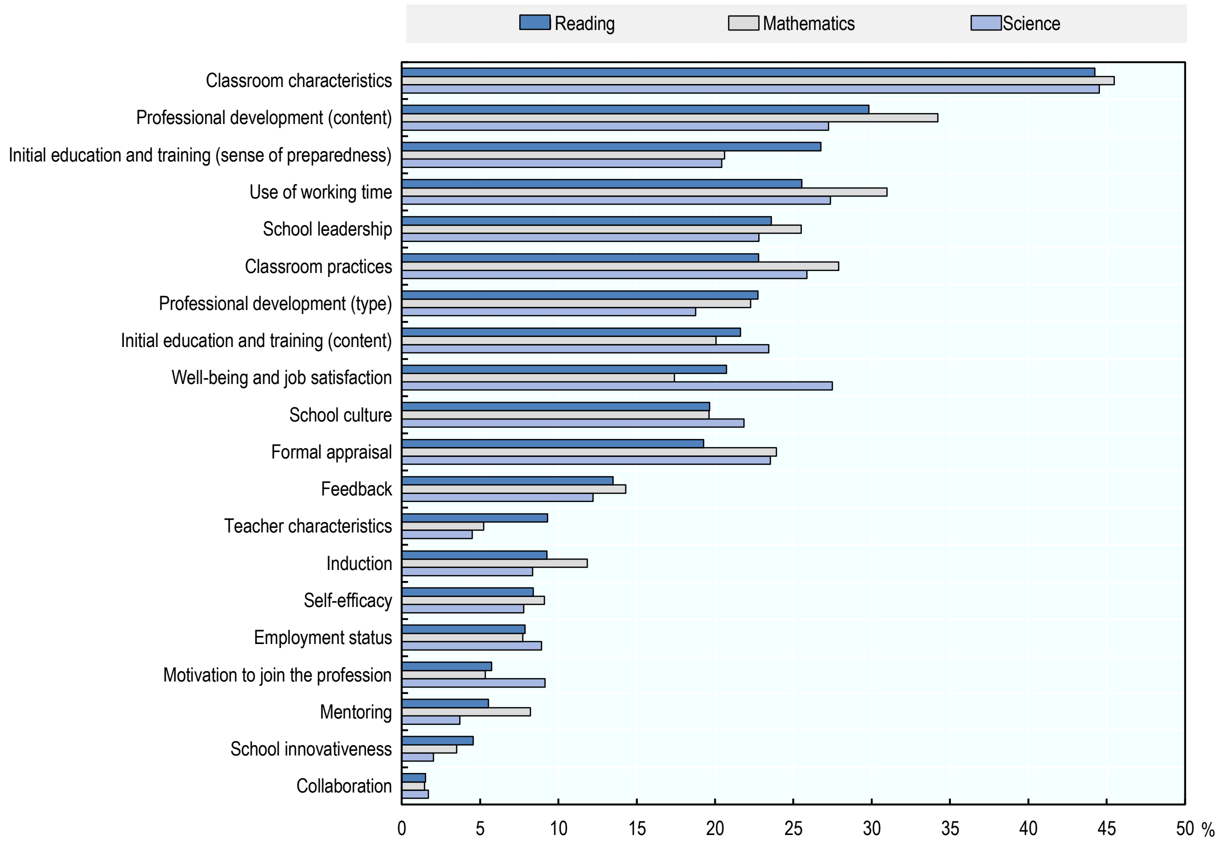
Do teacher and school factors matter equally across countries and subject domains?

The findings presented in the previous sections mainly focused on the teacher and school dimensions that matter for student performance for the overall population surveyed within the TALIS-PISA link in the case of the lasso regression analysis, and on average across countries in the case of the variance decomposition analysis. However, these results may mask important differences across the TALIS-PISA link countries and economies. Therefore, this section examines cross-country patterns. Student performance is regressed on each teacher and school dimension (taken separately) that is flagged by the lasso estimation results as an important predictor and explains a substantial part of between-school variance in student performance (i.e. 20% or above on average across subjects).

Besides revealing how the teacher and school factors that matter vary across countries and economies, the nested multiple linear regressions provide insights regarding the potential role of classmates' characteristics as confounding factors in the association between a teacher or school factor and student achievement. While controls for student characteristics such as gender, immigrant background and the index of economic, social and cultural status are always included in these regression models, the controls for the average classroom composition[20] within the school (also referred to as classmates' characteristics) are introduced in the augmented version of the models. Since these controls for classmates' characteristics are introduced in the nested multiple linear regression models while students' own characteristics are also included, they account for the mediating role of the academic abilities, behaviour problems and socio-economic and linguistic background of classmates.

Figure 2.4. School average differences in student achievement explained by teacher and school factors

Percentage of between-school variance in PISA scores, by dimension and by subject (TALIS-PISA link average)



Notes: Teacher variables are averaged for subject domain teachers. Thus, the analyses are based on samples restricted to schools with at least one subject domain teacher. The sum of between-school variances in student performance explained across all dimensions exceeds 100%, since such a sum does not take into account that the different dimensions are correlated with each other. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Values are ranked in descending order of the percentage of variance in PISA reading score explained at the school level.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.2, 2.3 and 2.4.

StatLink https://doi.org/10.1787/888934223726

Moreover, the country-level analysis does not focus solely on subject teachers, but it also looks at the relationships of interest by considering all teachers within the school. This complementary approach serves two goals: 1) it allows for revealing significant relationships that might be undetected due to low sample sizes given that the analyses are conducted at the country level; 2) it provides an opportunity to identify differential teacher and school effects on student outcomes (Box 2.2).

Based on lasso regression results presented above, six dimensions, including teachers' classroom practices, classroom characteristics, teachers' well-being and job satisfaction, teachers' use of working time, school culture and school leadership, have a significant association with student achievement in reading, mathematics and science. These dimensions also explain a significant part of the average school-level differences in student performances. Hence, this section examines how the relationships between the six teacher and school dimensions and student achievement in reading, mathematics and science vary across countries and subject domains. Moreover, this section also looks at the role of classmates' characteristics as a mediating factor in the relationship between specific teacher and school dimensions and student performance. It is also important to stress that, similarly to the lasso regression results, causal interpretation of the country-level regressions is not possible. All the results presented here are correlational and should be interpreted accordingly.

Box 2.2. How does the inclusion of non-subject teachers benefit the interpretation of the country-level findings?

Cross-country patterns in the relationships between teacher and school factors, and student performances are explored by considering all teachers within the school. As compared to analysing the overall population surveyed within the TALIS-PISA link, moving to country-level analysis leads to smaller sample sizes, which reduces the probability of identifying significant relationships. Therefore, to raise statistical power, the country-level analysis does not focus solely on subject teachers but also looks at relationships of interest by considering all teachers within the school.

If a teacher or a school factor matters for student achievement in a given subject by considering all teachers in the school, then it can be assumed that this factor also matters when focusing on teachers of that given subject. Subsequently, if a significant association is only established when all teachers are taken into account but not when the focus is on subject teachers, then it can be assumed that the lack of statistical power is at play.

On the other hand, a relationship may only be significant when focusing on subject teachers, but not when all teachers' practices or characteristics are taken into account. This suggests that the given factor matters specifically for subject teachers. Indeed, it can be assumed that certain teacher characteristics and practices are subject-specific and they can cancel each other out if averaged across teachers teaching different subjects. Past research shows that differential teacher and school effects on student outcomes may exist (Reynolds et al., 2014[29]; Rockoff, 2004[15]; Seidel and Shavelson, 2007[7]).

Thus, considering all teachers for the country-level analyses not only allows for revealing significant relationships that might be undetected due to low sample size, but it also provides an opportunity to identify differential teacher and school effects on student outcomes.

Sources: Reynolds, D. et al. (2014[29]), "Educational effectiveness research (EER): A state-of-the-art review", *School Effectiveness and School Improvement*, http://dx.doi.org/10.1080/09243453.2014.885450; Rockoff J. (2004[15]), "The impact of individual teachers on student achievement: Evidence from panel data", *American Economic Review*, http://dx.doi.org/10.1257/0002828041302244; Seidel, T. and R. Shavelson (2007[7]), "Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results", *Review of Educational Research*, http://dx.doi.org/10.3102/0034654307310317.

Teachers' classroom practices

The teacher dimension of classroom practices is consistently highlighted by the lasso regressions and the variance decomposition analysis as well as by education research as an important predictor of student performance. According to country-level regression results, as the average school teacher spends more class time on actual teaching and learning, students tend to perform better in all three subjects covered by PISA (i.e. reading, mathematics and science), in the Czech Republic, Turkey and on average across the countries and economies participating in the TALIS-PISA link (Figure 2.5). In Denmark, there is a positive relationship for mathematics and science, but for reading the association is not statistically significant.

When focusing on subject teachers, the relationship between student achievement and class time spent on actual teaching and learning by the school's average teacher becomes non-significant for reading and science in most countries (Figure 2.5). This may be an artefact of smaller sample sizes as the analyses are restricted to those schools that have at least one subject domain teacher. Yet, in Georgia, the relationship between student performance in science and the average class time spent on actual teaching and learning within the school becomes significant only when the focus is on science teachers. This suggests that, in the case of Georgia, the class time spent on actual teaching and learning by the school's average teacher may matter specifically for student performance in science.

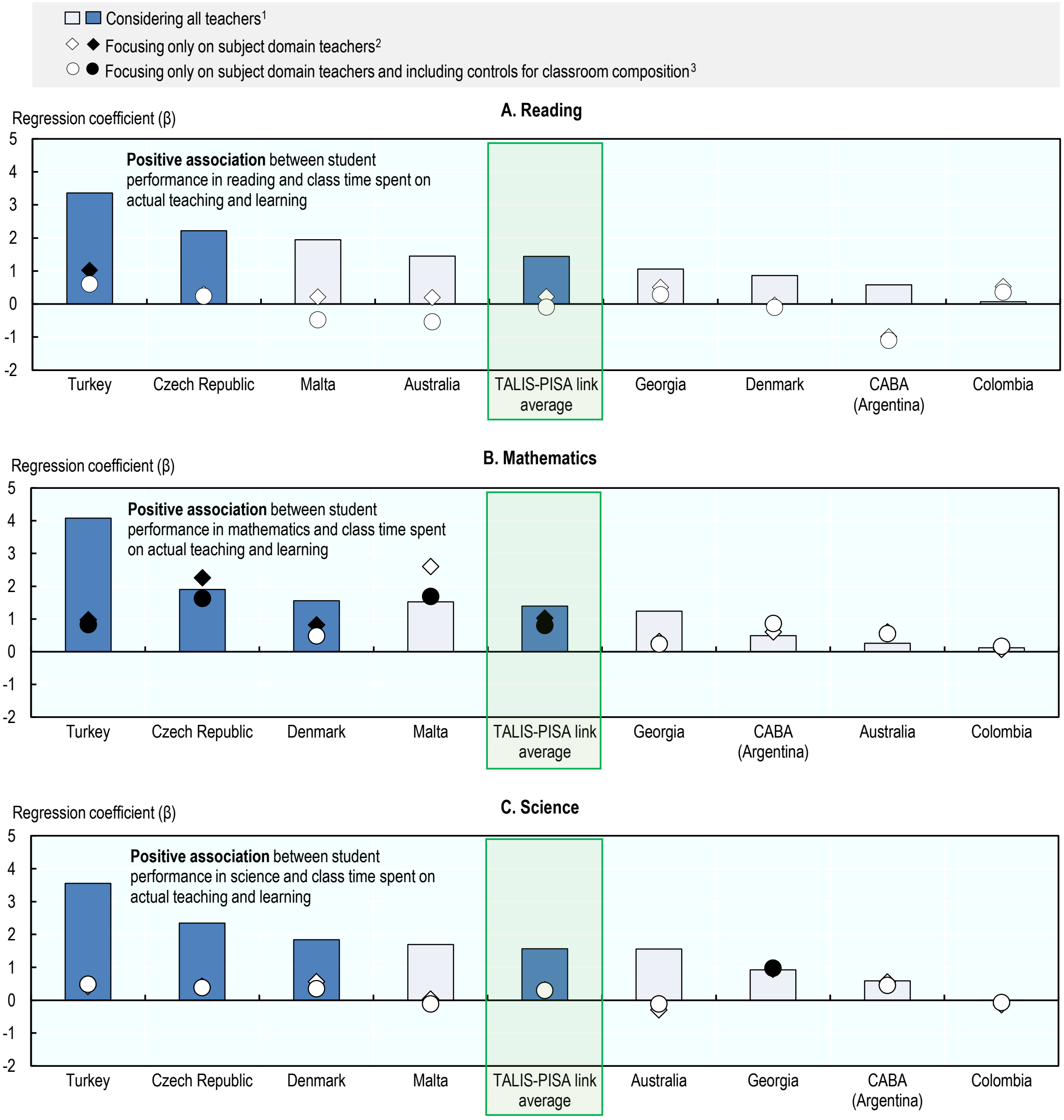
After accounting for classmates' characteristics, the association between student achievement and class time spent on actual teaching and learning by the school's average teacher remains significant only in the case of mathematics performance (Figure 2.5). This holds true in the Czech Republic, Turkey and on average across participating countries and economies. These results suggest that, especially in the case of reading and science classes, those with high academic achievers and students from privileged backgrounds likely have more exposure to actual teaching and learning than classes with a higher number of students who are struggling and come from socio-economically disadvantaged homes. These findings also indicate that, in general, the use of class time may be more closely linked to student performance in mathematics than in reading and science. This is also confirmed by the fact that classroom practices explain a larger share of the between-school variation in student performances in mathematics (28%) as compared to reading (23%), especially, and science (26%), to a lesser extent (Figure 2.4). Indeed, unlike reading, to which students are more exposed in all their academic activities as well as everyday life, mathematics and science are mainly learned at school (Reynolds et al., 2014[29]). In addition, there is research evidence for teacher and school effects being larger in mathematics and science than in reading (Reynolds et al., 2014[29]; Seidel and Shavelson, 2007[7]).

Classroom disciplinary climate, which can also be regarded as a partial measure for teachers' classroom management skills, is another element within classroom practices that turns out to be closely linked to student achievement. Indeed, based on country-level regression analysis, students who attend schools where there are classroom disciplinary issues tend to perform worse in all subjects in CABA (Argentina), the Czech Republic, Turkey and on average across participating countries and economies (Tables 2.5, 2.9 and 2.13).

The association between student achievement and school-level classroom disciplinary climate becomes non-significant in most countries and economies after controlling for classmates' characteristics (Tables 2.6, 2.8, 2.10, 2.12, 2.14 and 2.16). This suggests that students whose classmates are from privileged backgrounds might not only benefit from more class time spent on actual teaching and learning than their peers from socio-economically disadvantaged homes, but they also tend to attend schools where disciplinary problems are less of an issue.

Figure 2.5. Relationship between class time spent on teaching and student achievement

Change in PISA score associated with the average class time spent on actual teaching and learning at the school, by subject



1. Teacher variables are averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
3. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher. Besides controlling for student characteristics, the following controls are also included in order to account for the average classroom composition within the school: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees.

Notes: Results of linear regressions are based on responses of 15-year-old students and teachers. Controlling for the following classroom practices of teachers: teachers' autonomy over planning and teaching, perceived disciplinary climate, use of practices related to clarity of instruction, use of practices related to cognitive activation, use of assessment practices such as administering own assessment, providing written feedback on student work in addition to marking, letting students evaluate their own progress and observing students when working on particular tasks and providing immediate feedback; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

Countries and economies are ranked in descending order of the change in PISA score associated with the average class time spent on actual teaching and learning, considering all teachers at the school.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.5, 2.6, 2.8, 2.9, 2.10, 2.12, 2.13, 2.14 and 2.16.

StatLink https://doi.org/10.1787/888934223745

Teachers' well-being and job satisfaction

Teachers' well-being and job satisfaction is a dimension that is found to matter for student performance by the lasso regression and the variance decomposition analyses presented above. According to country-level regression results, the more teachers are satisfied with the work environment on average at the school the better students tend to perform in all three subjects covered by PISA. This relationship holds in Australia, the Czech Republic, Georgia, Turkey and on average across participating countries and economies (Figure 2.6).

However, after controlling for classmates' characteristics, the relationship is no longer significant for student performance in reading and mathematics across most countries and economies (Figure 2.6). This indicates that, in the case of reading and mathematics, the association between student performance and teachers' satisfaction with the work environment may largely be driven by classmates' characteristics.

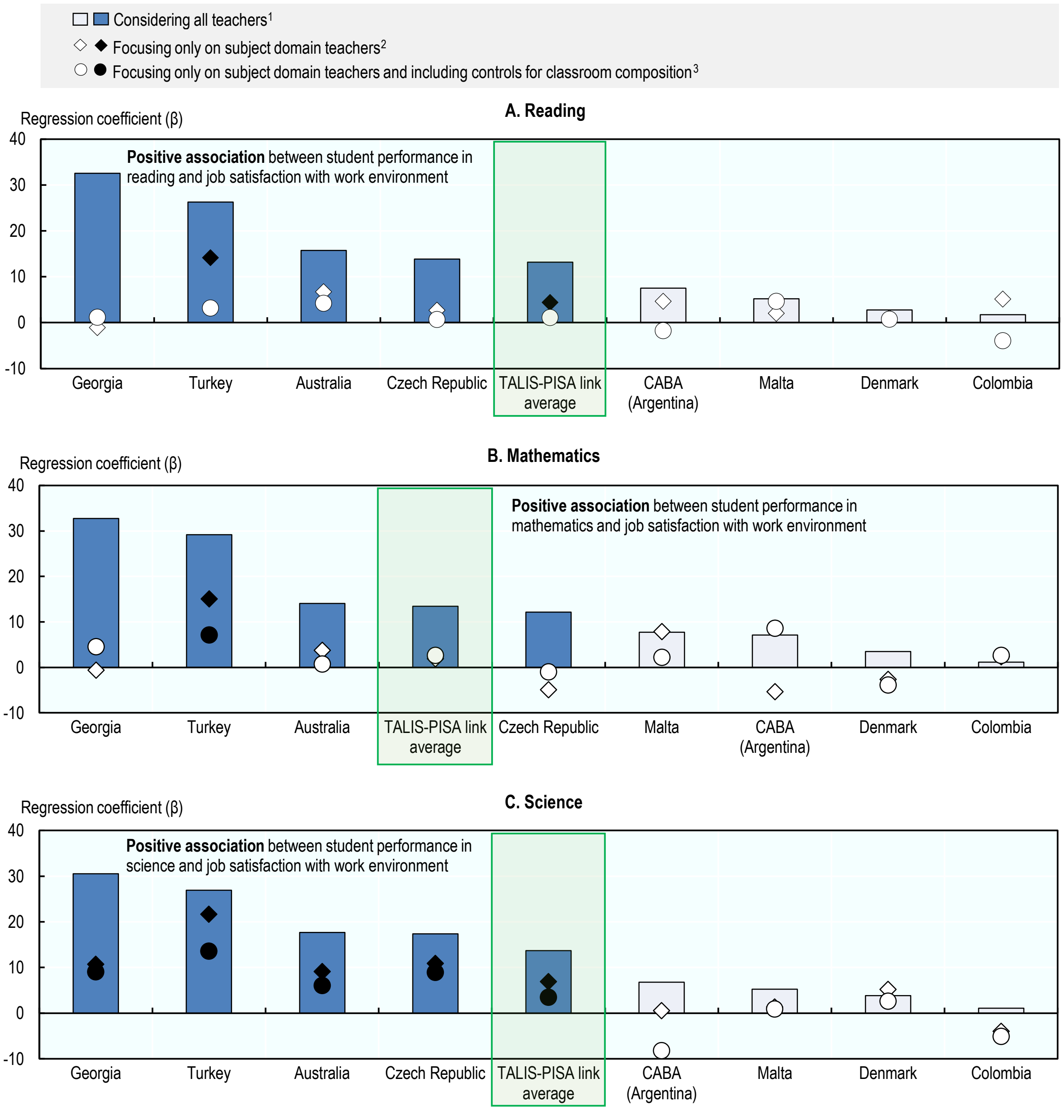
Teachers' satisfaction with the work environment seems to matter almost exclusively for student achievement in science once the average classroom composition within the school is accounted for and the focus shifts to subject teachers. This is the case in Australia, the Czech Republic, Georgia, Turkey and on average across participating countries and economies (Figure 2.6). Thus, country-level analysis suggests that teachers' satisfaction with the work environment may be more closely related to student performance in science than in reading and mathematics. This finding is in line with the results of the variance decomposition analysis presented in the previous section, which shows that the teacher well-being and job satisfaction dimension can explain a considerably larger share of the average school-level differences in performance in science (27%) than in reading (21%) and mathematics (17%) (Figure 2.4).

Similarly to why the use of class time may be most pertinent to mathematics, the close relationship between teachers' satisfaction with the work environment and student achievement in science can be also explained by the fact that, unlike reading, students mainly acquire their knowledge in science and mathematics at school (Reynolds et al., 2014[29]). Thus, science teachers' satisfaction with the work environment may indeed be more closely related with student achievement since students are most likely to learn about science in school. This can also explain why research points towards larger teacher and school effects in mathematics and science than in reading (Reynolds et al., 2014[29]; Seidel and Shavelson, 2007[7]). Moreover, work environment may matter specifically for science since teaching this subject has certain requirements regarding the work environment, such as a well-equipped school lab, that are critical for teachers to do their work properly.

Lasso regression results suggest that, besides teachers' satisfaction with the work environment, the extent to which workload is an important source of stress is also an important predictor of student achievement. In Colombia, the Czech Republic, Denmark and on average across participating countries and economies, students who attend schools where teachers, on average, report workload, including lesson preparation, number of lessons to teach, marking, administrative work and extra duties due to absent teachers, as an important source of stress, tend to perform better in all three subjects covered by PISA (Tables 2.25, 2.29 and 2.33). The relationship also becomes significant in CABA (Argentina) in the case of reading performance, when only reading teachers are included in the analysis (Table 2.26). This indicates that in CABA (Argentina) workload-induced stress may matter specifically for student achievement in reading. Overall, these findings may signal teachers' commitment and dedication to their work as well as point to a reverse causal relationship as highly competitive school environments (attended by higher-performing students) can lead to workload being an important source of stress for teachers.

Figure 2.6. Relationship between teachers' satisfaction with work environment and student achievement

Change in PISA score associated with the average job satisfaction with work environment at the school, by subject



1. Teacher variables averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
3. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher. Besides controlling for student characteristics, the following controls are also included in order to account for the average classroom composition within the school: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees.
Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following aspects of well-being and job satisfaction: workplace well-being and stress, workload stress, job satisfaction with profession, teachers' satisfaction with the salary, teachers' satisfaction with the terms of the teaching contract apart from salary (e.g. benefits, work schedule), teachers' views of the way different stakeholders value the profession; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam. Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in PISA score associated with the average job satisfaction with work environment, considering all teachers at the school.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.25, 2.26, 2.28, 2.29, 2.30, 2.32, 2.33, 2.34 and 2.36.

StatLink https://doi.org/10.1787/888934223764

The association between student performance in science and the extent to which science teachers report workload as an important source of stress on average within the school seems to be mainly driven by classmates' characteristics. The relationship in the case of science performance is no longer significant in any of the participating countries and economies once classmates' characteristics are taken into account (Table 2.36). Nevertheless, while focusing on subject teachers, the association remains significant in CABA (Argentina), the Czech Republic, Denmark and on average across participating countries and economies in the case of reading performance, and in Colombia, the Czech Republic and on average across participating countries and economies in the case of mathematics performance, even after accounting for classmates' characteristics (Tables 2.28 and 2.32).

Teachers' use of working time

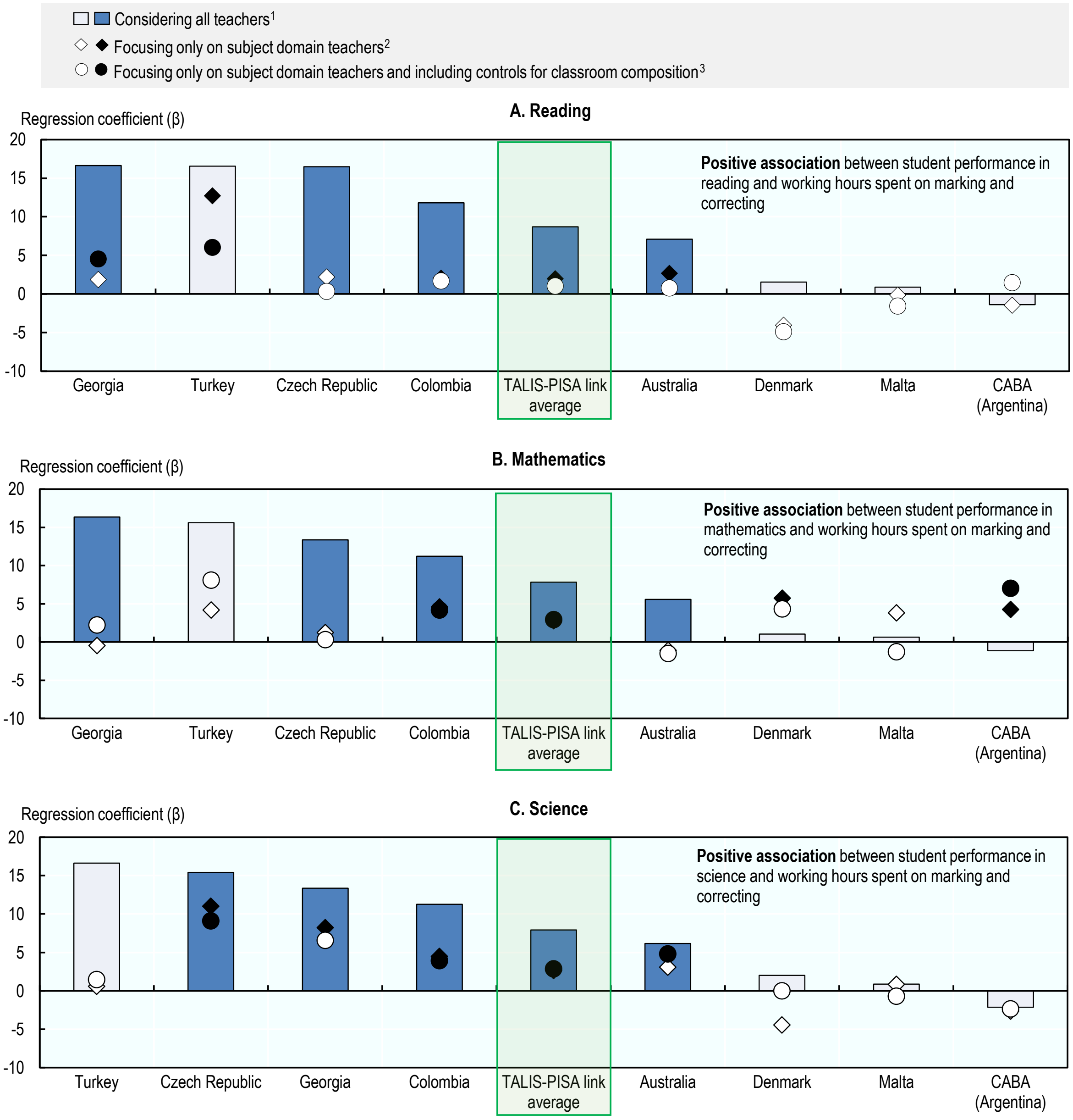
Teachers' use of working time is another teacher dimension with a potential indirect effect on student achievement that is flagged by the lasso regression and the variance decomposition analyses presented in previous sections. According to country-level regressions, as the amount of working hours teachers spend on marking and correcting increases on average at the school, the better students tend to perform in all three subjects covered by PISA in Australia, Colombia, the Czech Republic, Georgia and on average across participating countries and economies (Figure 2.7). This finding may point to a reverse causal relationship as highly competitive school environments (attended by higher-performing students) are characterised by more frequent feedback to students in the form of tests and exams.

As the focus shifts to the average working hours of subject teachers, the relationship also becomes significant for Turkey in the case of reading performance, and for CABA (Argentina) and Denmark in the case of mathematics performance (Figure 2.7). These results suggest that in Turkey the amount of working hours the school's average teacher spends on marking and correcting may matter specifically for student performance in reading, while in CABA (Argentina) and Denmark, marking and correcting may matter more for reading and mathematics compared to student performance in science.

Overall, the relationship between student achievement and teachers' average working hours spent on marking and correcting seems to be partly driven by the average classroom composition within the school. After controlling for classmates' characteristics, the association remains significant for student achievement in reading in the case of Georgia and Turkey, for performance in mathematics in the case of CABA (Argentina), Colombia and the TALIS-PISA link average, and for performance in science in the case of Australia, Colombia, the Czech Republic and the TALIS-PISA link average (Figure 2.7).

Figure 2.7. Relationship between time spent by teachers on marking and correcting student work and student achievement

Change in PISA score associated with the average working hours spent on marking and correcting at the school, by subject



1. Teacher variables are averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
3. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher. Besides controlling for student characteristics, the following controls are also included in order to account for the average classroom composition within the school: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees.

Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following elements of teachers' use of working time: total working hours, total teaching hours and teachers' use of working time on tasks other than marking and correcting (such as individual planning or preparation of lessons either at school or out of school, general administrative work, etc.); and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

Countries and economies are ranked in descending order of the change in PISA score associated with the average working hours spent on marking and correcting, considering all teachers at the school.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.37, 2.38, 2.40, 2.41, 2.42, 2.44, 2.45, 2.46 and 2.48.

StatLink https://doi.org/10.1787/888934223783

Classmates' characteristics

Regarding the characteristics of the classroom, one of the factors that shows the most consistent country-level pattern in student achievement across all three subjects is the concentration of students from socio-economically disadvantaged homes. In Australia, Colombia and Georgia, as the average concentration of students from socio-economically disadvantaged homes in the classrooms increases, the worse students tend to perform in all subjects covered by PISA (Figure 2.8). Since these findings hold while accounting for student's own socio-economic background, they point to the potential presence of peer effects. Indeed, a student's performance can be negatively affected by classmates with limited social, economic and cultural resources because this reflects the concentration of important disadvantages in the student's local community. Such concentration of disadvantages may also affect the student's cognitive and socio-emotional development due to fewer available material learning resources at the school and altered teaching strategies (in addition, see findings in Chapter 3).

As the focus shifts exclusively to subject teachers, the relationships become significant for Denmark in reading, for Malta in reading and mathematics, and for the TALIS-PISA link average for all three subjects (Figure 2.8). These findings indicate that peer effects may matter specifically in reading classes in Denmark and Malta, and in mathematics classes in Malta.

It is interesting to note that, based on TALIS-PISA link data, there are no indications for the presence of peer effects induced by the concentration of socio-economic disadvantage in CABA (Argentina), the Czech Republic and Turkey (Figure 2.8). In the case of CABA (Argentina) and the Czech Republic, this finding may be explained by the fact that disadvantaged students are often concentrated in schools with a small proportion of high achievers. Indeed, based on the PISA 2018 results, the probability that a typical disadvantaged student was enrolled in the same school as high achievers was less than one in eight in Argentina[21] and the Czech Republic (OECD, 2019[64]). Thus, detecting peer effects triggered by the concentration of disadvantage may be more challenging in these two education systems given that students with different ability levels and socio-economic status are less likely to attend the same school.

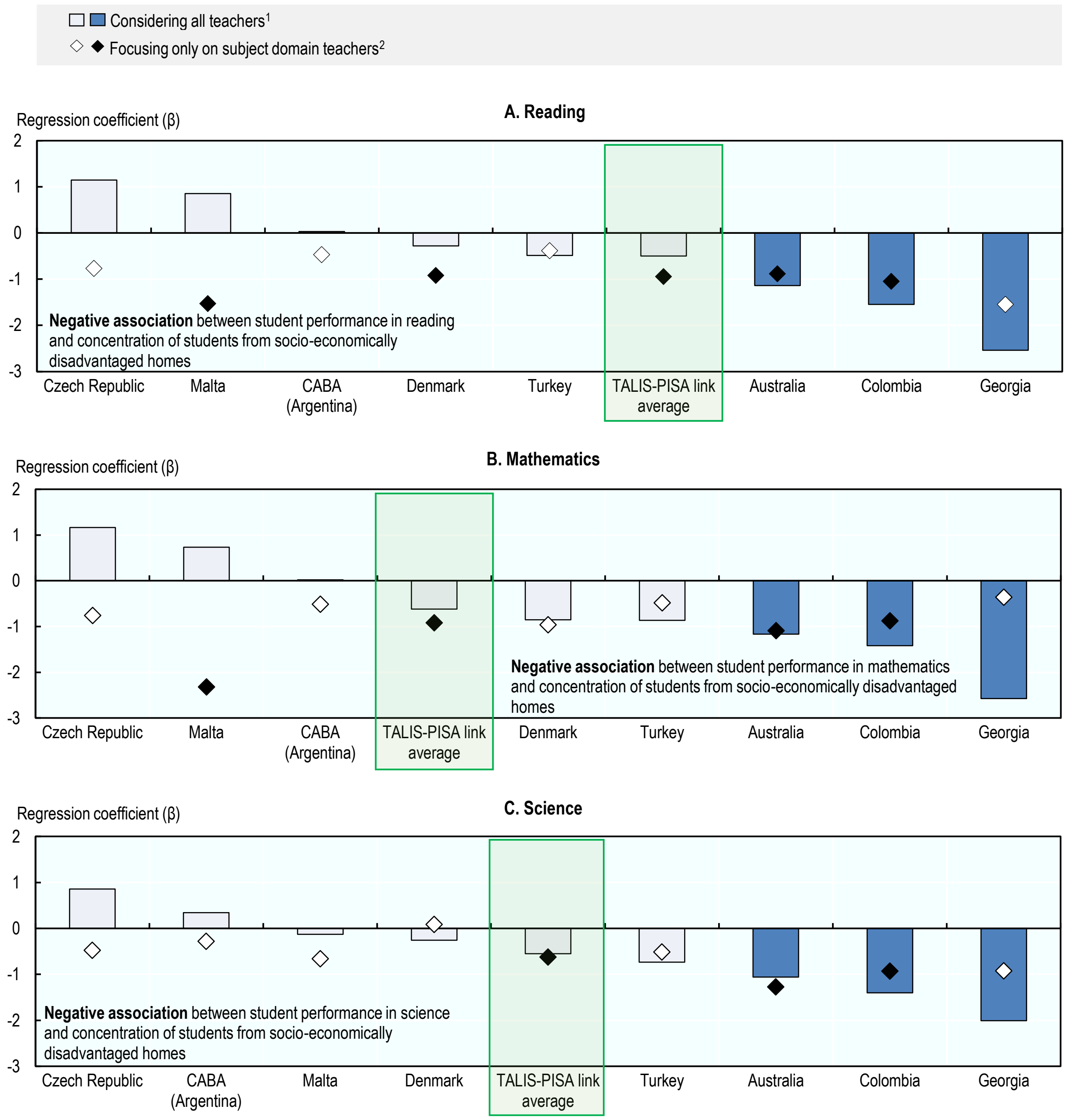
Another element within classmates' characteristics that shows a clear pattern across countries and economies is the share of academically gifted students. As the average concentration of academically gifted students in the classrooms increases, the better students tend to perform in all subjects covered by PISA, in Australia, CABA (Argentina), the Czech Republic, Turkey and on average across participating countries and economies (Figure 2.9). The only country where the association between the average concentration of academically gifted students in the classrooms and student achievement is not significant for any of the subject domains is Colombia.

These results can signal the presence of academic segregation. This means that academically gifted students, who fulfil their potential and become high achievers, are likely to attend schools where other students also tend to be high achievers. Indeed, based on the PISA 2018 results, there are education systems where low- and high-achieving students tend to be clustered in distinct schools (OECD, 2019[64]). For example, in the Czech Republic and Turkey, high achievers in reading are especially likely to be concentrated in certain schools.[22] By contrast, the degree of academic segregation, based on PISA reading scores, is low in Australia and Denmark. In these countries, students with different ability are likely to attend the same school.

As the relationship between the average concentration of academically gifted students in the classrooms and student achievement holds in all participating countries and economies, except for Colombia, irrespective of the degree of academic segregation, these results can also point to the potential presence of peer effects. Indeed, a student's performance can be positively affected by classmates with higher innate ability through an increase in motivation, competition and career aspirations (OECD, 2020[43]; Sacerdote, 2011[63]).

Figure 2.8. Relationship between concentration of students from socio-economically disadvantaged homes and student achievement

Change in PISA score associated with the average concentration of students from socio-economically disadvantaged homes at the school, by subject

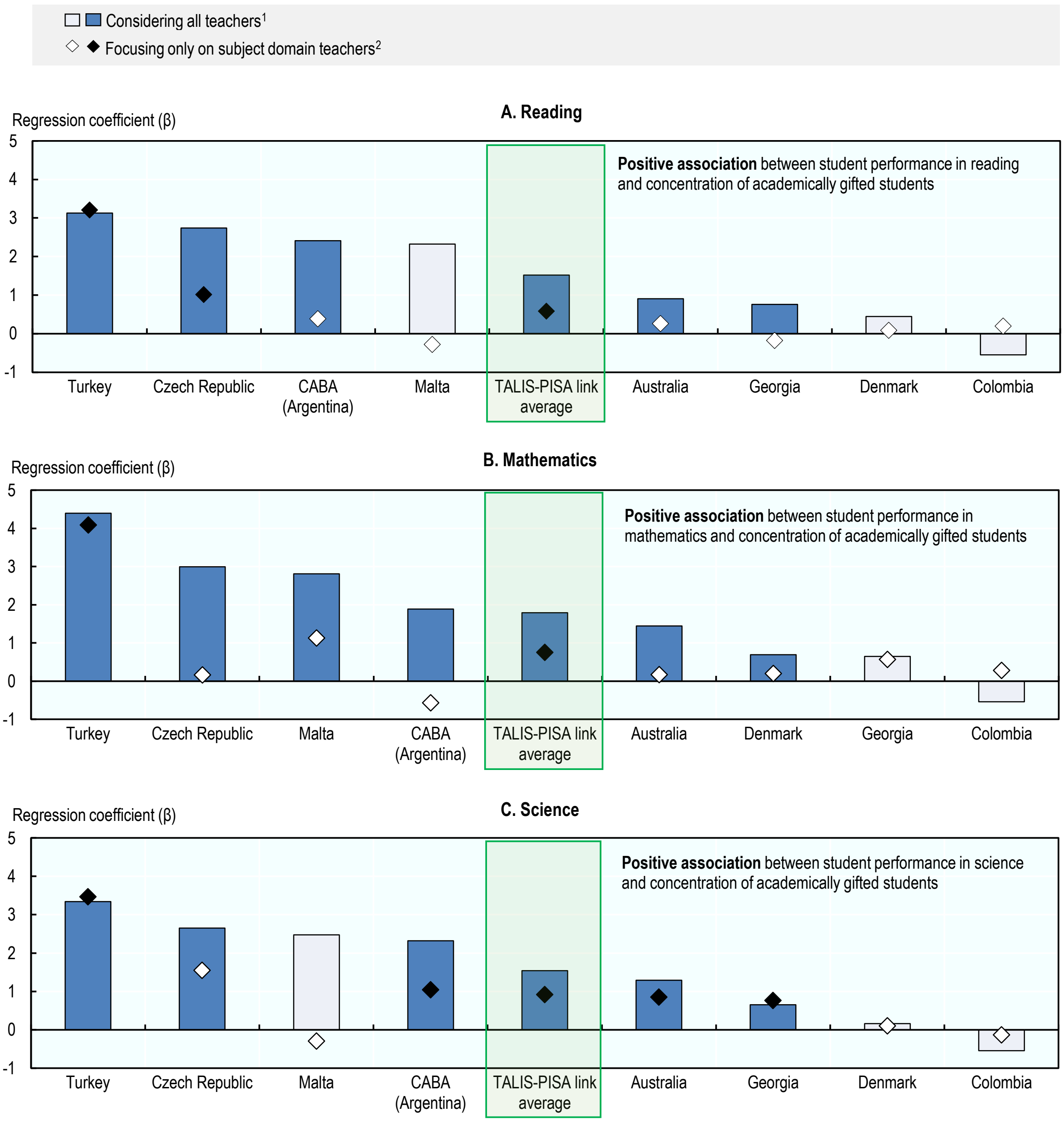


1. Teacher variables are averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, academically gifted students, students who are immigrants or with migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in PISA score associated with the average concentration of students from socio-economically disadvantaged homes, considering all teachers at the school.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.49, 2.50, 2.51, 2.52, 2.53 and 2.54.

StatLink https://doi.org/10.1787/888934223802

Figure 2.9. Relationship between concentration of academically gifted students and student achievement

Change in PISA score associated with the average concentration of academically gifted students at the school, by subject



1. Teacher variables are averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in PISA score associated with the average concentration of academically gifted students, considering all teachers at the school.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.49, 2.50, 2.51, 2.52, 2.53 and 2.54.

StatLink https://doi.org/10.1787/888934223821

School culture

Findings of lasso regression analysis presented above showed the importance of the involvement of parents and community in school-related activities for students' academic achievement. Similar results are found at the country level. In Australia, Colombia, Denmark and on average across participating countries and economies, students who attend schools where stakeholders (i.e. parents and community) are involved in school-related activities tend to perform better in the PISA test in all subjects (Figure 2.10).

However, after controlling for classmates' characteristics, the association between students' academic achievement and the involvement of parents and the community in school-related activities is no longer significant for most countries and economies participating in the TALIS-PISA link (Figure 2.10). The academic abilities, behaviour problems, and socio-economic and linguistic background of classmates seem to play an important mediating role in the association between student performance and stakeholder involvement in school-related activities. Yet, there are exceptions. In Denmark, the relationship still holds for all subjects even if the average classroom composition at the school is taken into account. The same holds for students' science performance in Colombia, where stakeholder involvement seems to matter for student achievement irrespective of classmates' characteristics.

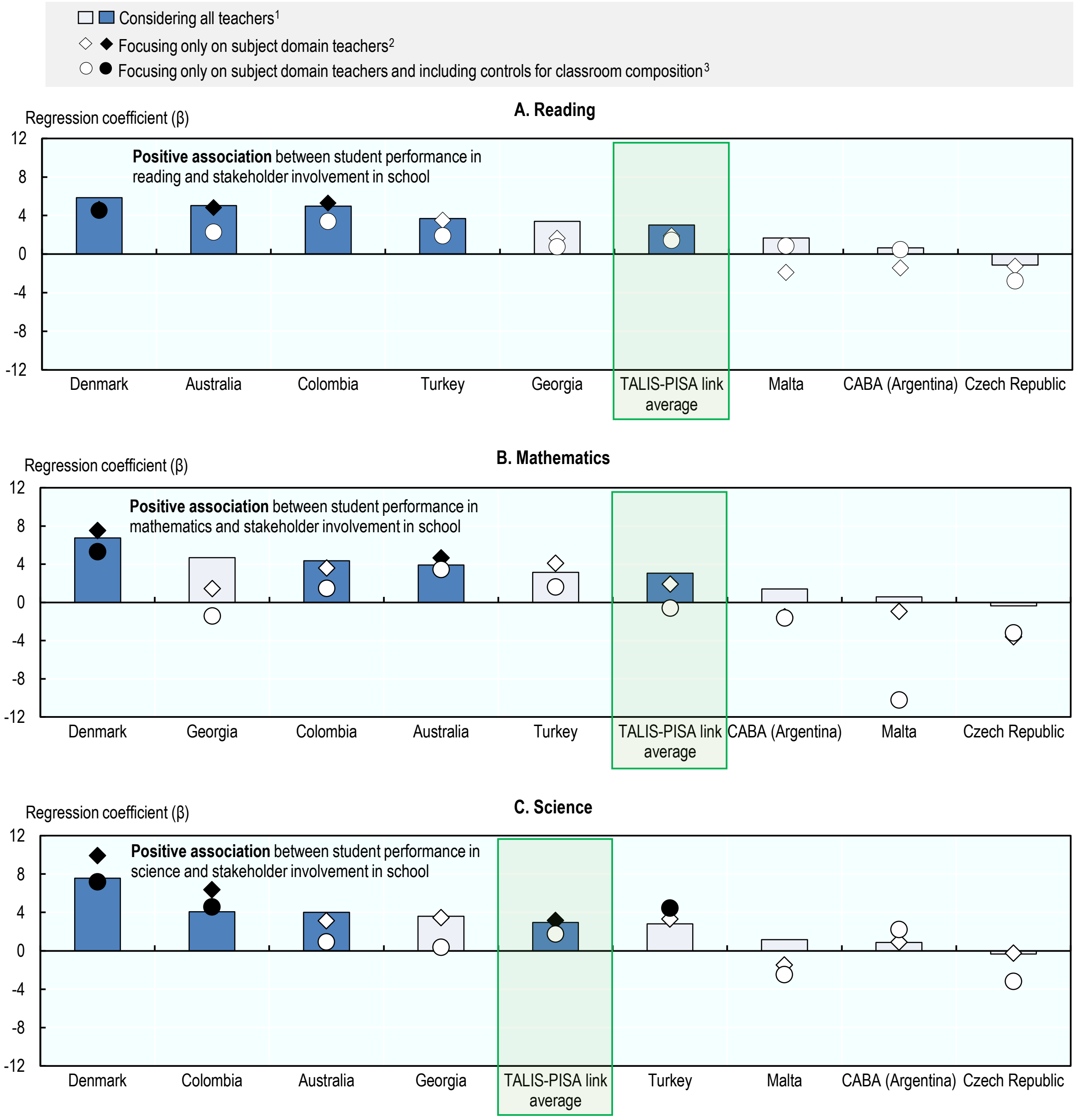
Teacher-student relations is another element of school culture highlighted by lasso regression analysis as being an important predictor of student achievement. In Australia, Turkey and on average across participating countries and economies, students who attend schools where teachers and students get along well also tend to perform well in all subjects covered by PISA (Tables 2.55, 2.59 and 2.63). The same relationship becomes significant in CABA (Argentina) when the focus is exclusively on mathematics teachers (Table 2.60). This suggests that, in the case of CABA (Argentina), teacher-student relations may matter specifically for student performance in mathematics. However, the relationship only remains significant in CABA (Argentina) in mathematics performance when classmates' characteristics are accounted for (Table 2.62). Hence, similarly to stakeholder involvement in school-related activities, classmates' characteristics have an important mediating effect in the relationship between teacher-student relations and student achievement.

School leadership

Based on the analyses presented in the previous sections, school leadership seems to matter for student performance. However, the country-level regression results indicate that the relationship of principals' leadership activities with student achievement mostly captures classmates' characteristics and classroom disciplinary issues. The only leadership activity for which there is a significant relationship in relation to student achievement for at least two participating countries and economies are principals' actions to solve disciplinary issues. Indeed, in CABA (Argentina), the Czech Republic and on average across participating countries and economies, students whose school leaders report having collaborated with teachers to solve classroom disciplinary problems in the 12 months prior to the survey tend to perform worse in all three subjects covered by PISA (Tables 2.91, 2.95 and 2.99). However, once classmates' characteristics are taken into account and the focus shifts to subject teachers, the relationship only remains significant in CABA (Argentina) in student performance in reading and mathematics, and in the Czech Republic in science performance (Tables 2.94, 2.98 and 2.102). This suggests that school leaders' collaboration with teachers to solve classroom disciplinary problems is in most cases closely related to the average classroom composition within the school.

Figure 2.10. Relationship between stakeholder involvement in school and student achievement

Change in PISA score associated with stakeholder (i.e. parents and local community) involvement in school, by subject



1. Teacher variables are averaged for all teachers within the school.
2. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher.
3. Since teacher variables are averaged only for subject domain teachers, the analysis is based on a sample restricted to schools with at least one subject domain teacher. Besides controlling for student characteristics, the following controls are also included in order to account for the average classroom composition within the school: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees.

Notes: Results of linear regressions based on responses of 15-year-old students and teachers. Controlling for the following aspects of school culture: collaborative school culture, teacher-student relations and teachers' actions towards achieving academic excellence; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in PISA score associated with stakeholder involvement in school, considering all teachers.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 2.55, 2.56, 2.58, 2.59, 2.60, 2.62, 2.63, 2.64 and 2.66.

StatLink https://doi.org/10.1787/888934223840

Summary

Drawing on the TALIS-PISA link data, this chapter identifies three main teacher and school factors that are not only found to matter for student achievement but are also within the reach of policy levers. These factors are: teachers' classroom practices (in particular the share of class time spent on actual teaching and learning), teachers' well-being and job satisfaction (in particular teachers' satisfaction with the work environment) and classroom characteristics (in particular the concentration of socio-economically disadvantaged students and the share of academically gifted students in the average classroom).

First, teachers' classroom practices seem to matter for student achievement. Average students tend to perform better the more class time school teachers spend on actual teaching and learning. This finding suggests that students' opportunity to learn, which is closely linked to the amount of time allocated to actual learning, is important for student achievement. When teachers do not actually teach and students do not learn in the class, it is usually either due to disciplinary issues or administrative tasks. More disruptive classrooms are more likely to have lower-achieving students, which, in turn, leads to more time spent on other activities such as keeping order or taking care of administrative tasks. Moreover, results suggest that the opportunities to learn are more closely linked to student performance in mathematics than in reading and science. This signals that the share of class time spent on actual teaching and learning matters specifically for those subjects that are mainly learnt at school, such as mathematics.

Second, teachers' job satisfaction seems to matter for student performance. The more satisfied teachers are with their work environment the better students tend to perform in school. This finding suggests that teachers' satisfaction with their work environment can play a role in teachers' attitudes, efforts and commitment, which, in turn, can lead to better performance. It can also signal the presence of self-enforcing dynamics. Teachers may be particularly satisfied with their school assignment when they work in schools attended by high-achieving students. In turn, these teachers might be particularly engaged in making their academically gifted students progress further. In addition, results signal the presence of differential effects across subjects as teachers' satisfaction with their work environment seems to be more closely related to student performance in science than in reading and mathematics. While this may be explained by the fact that, unlike reading, students mainly acquire their knowledge in science at school, it can also point to the fact that certain requirements regarding the work environment, such as a well-equipped school lab, are critical for science teachers to do their work properly.

Third, classroom characteristics, in particular, the concentration of socio-economic disadvantage and the concentration of academically gifted students, seem to matter for student achievement. The greater the average concentration of students from socio-economically disadvantaged homes in the classroom, the worse students tend to perform academically. Since this finding holds while accounting for students' own socio-economic background, it suggests the presence of peer effects. This finding suggests that a student's performance can be negatively affected if the student is surrounded by classmates with limited social, economic and cultural resources as this may reflect the concentration of significant disadvantages in the student's local community. Such concentration of disadvantages may also affect the student's cognitive development due to fewer available material learning resources at the school and altered teaching strategies. In fact, students, regardless of their own socio-economic background, tend to be advantaged scholastically if they attend a school whose students are from more advantaged socio-economic backgrounds. In addition, as the average concentration of academically gifted students in the classroom increases, students tend to perform better. This may signal the presence of academic segregation, as well as the presence of peer effects. Indeed, a student's performance can be also positively affected by classmates with higher innate ability through an increase in motivation, competition and career aspirations. Yet, high-performing students still tend to be less affected than their low-achieving peers by the composition of their classes. This indicates that addressing socio-economic and academic segregation of schools may be beneficial for both increasing student performance at the country level as well as improving equity.

References

Ainley, J. and R. Carstens (2018), “Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework”, *OECD Education Working Papers*, No. 187, OECD Publishing, Paris, https://dx.doi.org/10.1787/799337c2-en. [30]

Allen, R. and S. Sims (2017), *Improving Science Teacher Retention: Do National STEM Learning Network Professional Development Courses Keep Science Teachers in the Classroom?*, Wellcome Trust, London, https://www.stem.org.uk/resources/elibrary/resource/418071/improving-science-teacher-retention-do-national-stem-learning. [115]

Ammermueller, A. and J. Pischke (2009), “Peer effects in European primary schools: Evidence from the progress in international reading literacy study”, *Journal of Labor Economics*, Vol. 27/3, pp. 315-348, http://dx.doi.org/10.1086/603650. [120]

Angrist, J. and V. Lavy (2001), “Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools”, *Journal of Labor Economics*, Vol. 19/2, pp. 343-369, http://dx.doi.org/10.1086/319564. [88]

Angrist, J. and V. Lavy (1999), “Using Maimonides’rule to estimate the effect of class size on scholastic achievement”, *The Quarterly Journal of Economics*, Vol. 114/2, pp. 533-575, http://dx.doi.org/10.1162/003355399556061. [69]

Avanzi, L. et al. (2013), “Cross-validation of the Norwegian Teacher’s Self-Efficacy Scale (NTSES)”, *Teaching and Teacher Education*, Vol. 31, pp. 69-78, http://dx.doi.org/10.1016/J.TATE.2013.01.002. [51]

Avvisati, F. et al. (2014), “Getting parents involved: A field experiment in deprived schools”, *Review of Economic Studies*, Vol. 81/1, pp. 57-83, http://dx.doi.org/10.1093/restud/rdt027. [121]

Bakker, A. et al. (2007), “Job resources boost work engagement, particularly when job demands are high”, *Journal of Educational Psychology*, Vol. 99/2, pp. 274-284, http://dx.doi.org/10.1037/0022-0663.99.2.274. [58]

Barber, M. and M. Mourshed (2009), “Shaping the Future: How Good Education Systems Can Become Great in the Decade Ahead”, in *Report on the International Education Roundtable, 7 July 2009, Singapore*. [2]

Bascia, N. and C. Rottmann (2011), “What’s so important about teachers’ working conditions? The fatal flaw in North American educational reform”, *Journal of Education Policy*, Vol. 26/6, pp. 787-802, http://dx.doi.org/10.1080/02680939.2010.543156. [98]

Battistich, V. et al. (1997), “Caring school communities”, *Educational Psychologist*, Vol. 32/3, pp. 137-151, http://dx.doi.org/10.1207/s15326985ep3203_1. [73]

Baumert, J. et al. (2010), “Teachers’ mathematical knowledge, cognitive activation in the classroom, and student progress”, *American Educational Research Journal*, Vol. 47/1, pp. 133-180, http://dx.doi.org/10.3102/0002831209345157. [40]

Blazar, D. and M. Kraft (2017), “Teacher and teaching effects on students’ attitudes and behaviors”, *Educational Evaluation and Policy Analysis*, Vol. 39/1, pp. 146-170, http://dx.doi.org/10.3102/0162373716670260. [16]

Blömeke, S., J. Gustafsson and R. Shavelson (2015), "Beyond dichotomies: Competence viewed as a continuum", *Zeitschrift für Psychologie*, Vol. 223/1, pp. 3-13, http://dx.doi.org/10.1027/2151-2604/a000194. [39]

Boeskens, L. and D. Nusche (2021), "Not enough hours in the day: Policies that shape teachers' use of time", *OECD Education Working Papers*, No. 245, OECD Publishing, Paris, http://dx.doi.org/10.1787/15990b42-en. [55]

Borko, H. (2004), "Professional development and teacher learning: Mapping the terrain", *Educational Researcher*, Vol. 33/8, pp. 3-15, http://dx.doi.org/10.3102/0013189X033008003. [85]

Bouguen, A., J. Grenet and M. Gurgand (2017), "Does class size influence student achievement?", *IPP Policy Brief*, No. 28, Institut des Politiques Publiques, Paris, https://www.ipp.eu/wp-content/uploads/2018/07/n28-noteIPP-Sept2017.pdf. [71]

Bryk, A. and B. Schneider (2002), *Trust in Schools: A Core Resource for Improvement*, Russell Sage Foundation, New York, NY. [74]

Burke, M. and T. Sass (2013), "Classroom peer effects and student achievement", *Journal of Labor Economics*, Vol. 31/1, pp. 51-82, http://dx.doi.org/10.1086/666653. [61]

Caprara, G. et al. (2003), "Efficacy beliefs as determinants of teachers' job satisfaction", *Journal of Educational Psychology*, Vol. 95/4, pp. 821-832, http://dx.doi.org/10.1037/0022-0663.95.4.821. [45]

Carbonneau, N. et al. (2008), "The role of passion for teaching in intrapersonal and interpersonal outcomes", *Journal of Educational Psychology*, Vol. 100/4, pp. 977-987, http://dx.doi.org/10.1037/a0012545. [114]

Chapman, C. et al. (eds.) (2015), *Routledge International Handbook of Educational Effectiveness and Improvement Research: Research, Policy, and Practice*, Routledge, Abingdon; New York, NY. [102]

Chesnut, S. and H. Burley (2015), "Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis", *Educational Research Review*, Vol. 15, pp. 1-16, http://dx.doi.org/10.1016/j.edurev.2015.02.001. [52]

Chetty, R., J. Friedman and J. Rockoff (2014), "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates", *American Economic Review*, Vol. 104/9, pp. 2593-2632, http://dx.doi.org/10.1257/aer.104.9.2593. [9]

Chetty, R., J. Friedman and J. Rockoff (2014), "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood", *American Economic Review*, Vol. 104/9, pp. 2633-2679, http://dx.doi.org/10.1257/aer.104.9.2633. [10]

Chetty, R., N. Hendren and L. Katz (2016), "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment", *American Economic Review*, Vol. 106/4, pp. 855-902, http://dx.doi.org/10.1257/aer.20150572. [122]

Clotfelter, C. et al. (2008), "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina", *Journal of Public Economics*, Vol. 92/5-6, pp. 1352-1370, http://dx.doi.org/10.1016/j.jpubeco.2007.07.003. [116]

Cohen, J. et al. (2009), “School climate: Research, policy, practice, and teacher education”, *Teachers College Record*, Vol. 111/1, pp. 180-213. [75]

Collie, R., J. Shapka and N. Perry (2012), “School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy”, *Journal of Educational Psychology*, Vol. 104/4, pp. 1189-1204, http://dx.doi.org/10.1037/a0029356. [49]

Creemers, B. and L. Kyriakides (2015), “Process-product research: A cornerstone in educational effectiveness research”, *Journal of Classroom Interaction*, Vol. 50/2, pp. 107-119. [31]

Creemers, B. and L. Kyriakides (2010), “School factors explaining achievement on cognitive and affective outcomes: Establishing a dynamic model of educational effectiveness”, *Scandinavian Journal of Educational Research*, Vol. 54/3, pp. 263-294, http://dx.doi.org/10.1080/00313831003764529. [26]

Creemers, B. and L. Kyriakides (2007), *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*, Routledge, Abingdon, https://lib.ugent.be/nl/catalog/rug01:001240853. [32]

Darling-Hammond, L. (2017), “Teacher education around the world: What can we learn from international practice?”, *European Journal of Teacher Education*, Vol. 40/3, pp. 291-309, http://dx.doi.org/10.1080/02619768.2017.1315399. [3]

de Cuyper, N., H. de Witte and H. Van Emmerik (2011), “Temporary employment: Costs and benefits for (the careers of); employees and organizations”, *Career Development International*, Vol. 16/2, pp. 104-113, http://dx.doi.org/10.1108/13620431111115587. [100]

Dobbie, W. (2011), *Teacher Characteristics and Student Achievement: Evidence from Teach For America*, Harvard University, Cambridge, MA, https://www.hks.harvard.edu/publications/teacher-characteristics-and-student-achievement-evidence-teach-america. [21]

Dobbie, W. and R. Fryer (2013), “Getting beneath the veil of effective schools: Evidence from New York City”, *American Economic Journal: Applied Economics*, Vol. 5/4, pp. 28-60, http://dx.doi.org/10.1257/app.5.4.28. [28]

Dolton, P. et al. (2018), *Global Teacher Status: Index 2018*, Varkey Foundation, London, https://www.varkeyfoundation.org/media/4867/gts-index-13-11-2018.pdf. [118]

Duflo, E., P. Dupas and M. Kremer (2011), “Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya”, *American Economic Review*, Vol. 101/5, pp. 1739-1774, http://dx.doi.org/10.1257/aer.101.5.1739. [123]

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. [35]

Engel, L., D. Rutkowski and L. Rutkowski (2009), “The harsher side of globalisation: Violent conflict and academic achievement”, *Globalisation, Societies and Education*, Vol. 7/4, pp. 433-456, http://dx.doi.org/10.1080/14767720903412242. [76]

Feng, L. and T. Sass (2018), “The impact of incentives to recruit and retain teachers in ‘hard-to-staff’ subjects”, *Journal of Policy Analysis and Management*, Vol. 37/1, pp. 112-135, http://dx.doi.org/10.1002/pam.22037. [117]

Fischer, C. et al. (2018), "Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform", *Teaching and Teacher Education*, Vol. 72, pp. 107-121, http://dx.doi.org/10.1016/J.TATE.2018.02.011. [87]

Garet, M. et al. (2016), *Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher Professional Development (NCEE 2016-4010)*, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, DC, https://files.eric.ed.gov/fulltext/ED569154.pdf. [86]

Goddard, R. et al. (2015), "A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning", *American Journal of Education*, Vol. 121/4, pp. 501-530, http://dx.doi.org/10.1086/681925. [109]

Goddard, Y., R. Goddard and M. Tschannen-Moran (2007), "A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools", *Teachers College Record*, Vol. 109/4, pp. 877-896, https://www.tcrecord.org/content.asp?contentid=12871. [93]

Gomendio, M. (2017), *Empowering and Enabling Teachers to Improve Equity and Outcomes for All*, International Summit on the Teaching Profession, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264273238-en. [99]

Gordon, R., T. Kane and D. Staiger (2008), "Identifying effective teachers using performance on the job", in Furman, J. and J. Bordof (eds.), *Path to Prosperity: Hamilton Project Ideas on Income Security, Education, and Taxes*, Brookings Institution Press, Washington, DC, https://www.jstor.org/stable/10.7864/j.ctt1262d6. [19]

Hakanen, J., A. Bakker and W. Schaufeli (2006), "Burnout and work engagement among teachers", *Journal of School Psychology*, Vol. 43/6, pp. 495-513, http://dx.doi.org/10.1016/j.jsp.2005.11.001. [59]

Hallinger, P. (2018), "Bringing context out of the shadows of leadership", *Educational Management Administration and Leadership*, Vol. 46/1, pp. 5-24, http://dx.doi.org/10.1177/1741143216670652. [103]

Hallinger, P. (2015), "The evolution of instructional leadership", in *Assessing Instructional Leadership with the Principal Instructional Management Rating Scale*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-15533-3_1. [107]

Hallinger, P. (2011), "Leadership for learning: Lessons from 40 years of empirical research", *Journal of Educational Administration*, Vol. 49/2, pp. 125-142, http://dx.doi.org/10.1108/09578231111116699. [104]

Hallinger, P. and R. Heck (2010), "Leadership for learning: Does collaborative leadership make a difference in school improvement?", *Educational Management Administration & Leadership*, Vol. 38/6, pp. 654-678, http://dx.doi.org/10.1177/1741143210379060. [108]

Hanushek, E. (2006), "School resources", in Hanushek, E. and F. Welch (eds.), *Handbook of the Economics of Education*, Elsevier, http://dx.doi.org/10.1016/S1574-0692(06)02014-9. [65]

Hanushek, E. (2003), "The failure of input-based schooling policies", *The Economic Journal*, Vol. 113/485, pp. F64-F98, http://dx.doi.org/10.1111/1468-0297.00099. [66]

Hanushek, E. (1999), “Some findings from an independent investigation of the Tennessee STAR Experiment and from other investigations of class size effects”, *Educational Evaluation and Policy Analysis*, Vol. 21/2, pp. 143-163, http://dx.doi.org/10.3102/01623737021002143. [67]

Hanushek, E. and S. Rivkin (2010), “Generalizations about using value-added measures of teacher quality”, *American Economic Review*, Vol. 100/2, pp. 267-271, http://dx.doi.org/10.1257/aer.100.2.267. [11]

Hargreaves, A. (1992), “Time and teachers’ work: An analysis of the intensification thesis”, *Teachers College Record*, Vol. 94/1, pp. 87-108, https://www.tcrecord.org/content.asp?contentid=179. [56]

Hastie, T., R. Tibshirani and J. Friedman (2017), “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”, in *Springer Series in Statistics*, Springer, New York, https://web.stanford.edu/~hastie/ElemStatLearn//printings/ESLII_print12.pdf. [110]

Hastie, T., R. Tibshirani and M. Wainwright (2015), “Statistical learning with sparsity: The lasso and generalizations”*, Monographs on Statistics and Applied Probability*, No. 143, CRC Press, Boca Raton, FL, https://web.stanford.edu/~hastie/StatLearnSparsity_files/SLS.pdf. [111]

Hattie, J. (2009), *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*, Routledge, London. [5]

Hill, H., B. Rowan and D. Ball (2005), “Effects of teachers’ mathematical knowledge for teaching on student achievement”, *American Educational Research Journal*, Vol. 42/2, pp. 371-406, https://doi.org/10.3102/00028312042002371. [41]

Holzberger, D., A. Philipp and M. Kunter (2013), “How teachers’ self-efficacy is related to instructional quality: A longitudinal analysis”, *Journal of Educational Psychology*, Vol. 105/3, pp. 774-786, http://dx.doi.org/10.1037/a0032198. [50]

Hoxby, C. (2000), “The effects of class size on student achievement: New evidence from population variation”, *The Quarterly Journal of Economics*, Vol. 115/4, pp. 1239-1285, http://dx.doi.org/10.1162/003355300555060. [68]

Hoy, W., C. Tarter and A. Hoy (2006), “Academic optimism of schools: A force for student achievement”, *American Educational Research Journal*, Vol. 43/3, pp. 425-446, https://doi.org/10.3102/00028312043003425. [77]

Hoy, W. and A. Woolfolk (1993), “Teachers’ sense of efficacy and the organizational health of schools”, *The Elementary School Journal*, Vol. 93/4, pp. 355-372, https://doi.org/10.1086/461729. [80]

Ingersoll, R. and M. Strong (2011), “The impact of induction and mentoring programs for beginning teachers: A critical review of the research”, *Review of Educational Research*, Vol. 81/2, pp. 201-233, http://dx.doi.org/10.3102/0034654311403323. [83]

Ingvarson, L., M. Meiers and A. Beavis (2005), “Factors affecting the impact of professional development programs on teachers’ knowledge, practice, student outcomes and efficacy”, *Education Policy Analysis Archives*, Vol. 13/10, pp. 1-28, http://dx.doi.org/10.14507/epaa.v13n10.2005. [90]

Isoré, M. (2009), “Teacher Evaluation: Current Practices in OECD Countries and a Literature Review”, *OECD Education Working Papers*, No. 23, OECD Publishing, Paris, https://dx.doi.org/10.1787/223283631428. [101]

Jackson, C. (2018), “What do test scores miss? The importance of teacher effects on non-test score outcomes”, *Journal of Political Economy*, Vol. 126/5, pp. 2072-2107, http://dx.doi.org/10.1086/699018. [17]

Jackson, C., J. Rockoff and D. Staiger (2014), “Teacher effects and teacher-related policies”, *Annual Review of Economics*, Vol. 6, pp. 801-825, http://dx.doi.org/10.1146/annurev-economics-080213-040845. [12]

Jensen, B. and J. Reichl (2011), “Better teacher appraisal and feedback: Improving performance”, *Grattan Institute Report*, No. 2011-3 APR 2011, Grattan Institute, Melbourne, https://grattan.edu.au/wp-content/uploads/2014/04/081_report_teacher_appraisal.pdf. [91]

Jepsen, C. and S. Rivkin (2009), “Class size reduction and student achievement”, *Journal of Human Resources*, Vol. 44/1, pp. 223-250, http://dx.doi.org/10.3368/jhr.44.1.223. [72]

Kane, T., J. Rockoff and D. Staiger (2008), “What does certification tell us about teacher effectiveness? Evidence from New York City”, *Economics of Education Review*, Vol. 27/6, pp. 615-631, http://dx.doi.org/10.1016/J.ECONEDUREV.2007.05.005. [20]

Kane, T. and D. Staiger (2008), “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation”, *NBER Working Paper Series*, No. 14607, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w14607. [13]

Kersting, N. et al. (2012), “Measuring usable knowledge: Teachers’ analyses of mathematics classroom videos predict teaching quality and student learning”, *American Educational Research Journal*, Vol. 49/3, pp. 568-589, http://dx.doi.org/10.3102/0002831212437853. [42]

Klassen, R. and M. Chiu (2010), “Effects on teachers’ self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress”, *Journal of Educational Psychology*, Vol. 102/3, pp. 741-756, http://dx.doi.org/10.1037/a0019237. [60]

Klassen, R. et al. (2013), “Preservice teachers’ work stress, self-efficacy, and occupational commitment in four countries”, *European Journal of Psychology of Education*, Vol. 28/4, pp. 1289-1309, http://dx.doi.org/10.1007/s10212-012-0166-x. [46]

Klusmann, U. et al. (2008), “Teachers’ occupational well-being and quality of instruction: The important role of self-regulatory patterns”, *Journal of Educational Psychology*, Vol. 100/3, pp. 702-715, http://dx.doi.org/10.1037/0022-0663.100.3.702. [53]

Kraft, M. (2019), “Teacher effects on complex cognitive skills and social-emotional competencies”, *Journal of Human Resources*, Vol. 54/1, pp. 1-36, http://dx.doi.org/10.3368/JHR.54.1.0916.8265R3. [18]

Krueger, A. (2003), “Economic considerations and class size”, *The Economic Journal*, Vol. 113/485, pp. F34-F63, http://dx.doi.org/10.1111/1468-0297.00098. [70]

Kyriakides, L., C. Christoforou and C. Charalambous (2013), “What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching”, *Teacher and Teacher Education*, Vol. 36, pp. 143-152, http://dx.doi.org/10.1016/j.tate.2013.07.010. [33]

Lavy, V., O. Silva and F. Weinhardt (2012), "The good, the bad, and the average: Evidence on ability peer effects in schools", *Journal of Labor Economics*, Vol. 30/2, pp. 367-414, http://dx.doi.org/10.1086/663592. [62]

Le Donné, N., P. Fraser and G. Bousquet (2016), "Teaching Strategies for Instructional Quality: Insights from the TALIS-PISA Link Data", *OECD Education Working Papers*, No. 148, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jln1hlsr0lr-en. [36]

Leithwood, K., L. Leonard and L. Sharratt (1998), "Conditions fostering organizational learning in schools", *Educational Administration Quarterly*, Vol. 34/2, pp. 243-276, http://dx.doi.org/10.1177/0013161X98034002005. [95]

Marks, H. and S. Printy (2003), "Principal leadership and school performance: An integration of transformational and instructional leadership", *Educational Administration Quarterly*, Vol. 39/3, pp. 370-397, http://dx.doi.org/10.1177/0013161X03253412. [96]

Martin, M. et al. (2013), "Effective schools in reading, mathematics, and science at the fourth grade", in Martin, M. and I. Mullis (eds.), *TIMSS and PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade - Implications for Early Learning*, TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA), Chestnut Hill, MA, https://timssandpirls.bc.edu/timsspirls2011/downloads/TP11_Chapter_3.pdf. [78]

Mostafa, T. and J. Pál (2018), "Science teachers' satisfaction: Evidence from the PISA 2015 teacher survey", *OECD Education Working Papers*, No. 168, OECD Publishing, Paris, https://dx.doi.org/10.1787/1ecdb4e3-en. [127]

Muijs, D. (2011), "Researching leadership: Towards a new paradigm", in Townsend T. and MacBeath J. (eds.), *International Handbook of Leadership for Learning*, Springer, Dordrecht, http://dx.doi.org/10.1007/978-94-007-1350-5_9. [105]

Muijs, D. et al. (2014), "State of the art: Teacher effectiveness and professional learning", *School Effectiveness and School Improvement*, Vol. 25/2, pp. 231-256, http://dx.doi.org/10.1080/09243453.2014.885451. [23]

Muijs, D. and D. Reynolds (2001), *Effective Teaching: Evidence and Practice*, Sage Publications, London. [119]

Nilsen, T. and J. Gustafsson (2014), "School emphasis on academic success: Exploring changes in science performance in Norway between 2007 and 2011 employing two-level SEM", *Educational Research and Evaluation*, Vol. 20/4, pp. 308-327, http://dx.doi.org/10.1080/13803611.2014.941371. [79]

O'Donnell, R. and G. White (2005), "Within the accountability era: Principals' instructional leadership behaviors and student achievement", *NASSP Bulletin*, Vol. 89/645, pp. 56-71, http://dx.doi.org/10.1177/019263650508964505. [97]

OECD (2020), *Global Teaching InSights: A Video Study of Teaching*, OECD Publishing, Paris, https://dx.doi.org/10.1787/20d6f36b-en. [25]

OECD (2020), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/ca768d40-en. [37]

OECD (2020), *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/19cf08df-en. [43]

OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/5f07c754-en. [128]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b5fd1b8f-en. [64]

OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/1d0bc92a-en. [34]

OECD (2018), *Effective Teacher Policies: Insights from PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264301603-en. [4]

OECD (2014), *TALIS 2013 Results: An International Perspective on Teaching and Learning*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264196261-en. [44]

OECD (2013), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264201132-en. [124]

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. [92]

OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, Education and Training Policy, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264018044-en. [1]

Opfer, D. (2016), “Conditions and Practices Associated with Teacher Professional Development and Its Impact on Instruction in TALIS 2013”, *OECD Education Working Papers*, No. 138, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jlss4r0lrg5-en. [89]

Paniagua, A. and D. Istance (2018), *Teachers as Designers of Learning Environments: The Importance of Innovative Pedagogies*, Educational Research and Innovation, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264085374-en. [57]

Papay, J. and M. Kraft (2015), “Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement”, *Journal of Public Economics*, Vol. 130, pp. 105-119, http://dx.doi.org/10.1016/j. jpubeco.2015.02.008. [38]

Reeves, P., W. Pun and K. Chung (2017), “Influence of teacher collaboration on job satisfaction and student achievement”, *Teaching and Teacher Education*, Vol. 67, pp. 227-236, http://dx.doi.org/10.1016/j.tate.2017.06.016. [94]

Reynolds, D. and D. Muijs (2016), “Leading effective pedagogy”, in Harris, A. and M. Jones (eds.), *Leading Futures: Global Perspectives on Educational Leadership*, Sage Publications India, New Delhi. [106]

Reynolds, D. et al. (2014), “Educational effectiveness research (EER): A state-of-the-art review”, *School Effectiveness and School Improvement*, Vol. 25/2, pp. 197-230, http://dx.doi.org/10.1080/09243453.2014.885450. [29]

Rice, J. (2003), *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*, Economic Policy Institute, Washington, DC. [6]

Rivkin, S., E. Hanushek and J. Kain (2005), “Teachers, schools, and academic achievement”, *Econometrica: The Journal of the Econometric Society*, Vol. 73/2, pp. 417-458, https://doi.org/10.1111/j.1468-0262.2005.00584.x. [14]

Rockoff, J. (2008), “Does Mentoring Reduce Turnover and Improve Skills of New Employees? Evidence from Teachers in New York City”*, NBER Working Paper Series*, No. 13868, National Bureau of Economic Research, Cambridge, MA, http://www.nber.org/papers/w13868. [84]

Rockoff, J. (2004), “The impact of individual teachers on student achievement: Evidence from panel data”, *American Economic Review*, Vol. 94/2, pp. 247-252, http://dx.doi.org/10.1257/0002828041302244. [15]

Rockoff, J. et al. (2011), “Can you recognize an effective teacher when you recruit one?”, *Education Finance and Policy*, Vol. 6/1, pp. 43-74, http://dx.doi.org/10.1162/EDFP_a_00022. [22]

Sacerdote, B. (2011), “Peer effects in education: How might they work, how big are they and how much do we know thus far?”, in Hanushek, E., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1. [63]

Scheerens, J. (2001), “School effectiveness research”, in Smelser, N. and P. Baltes (eds.), *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/b0-08-043076-7/02438-4. [27]

Schmidt, W., P. Zoido and L. Cogan (2014), “Schooling Matters: Opportunity to Learn in PISA 2012”*, OECD Education Working Papers*, No. 95, OECD Publishing, Paris, https://dx.doi.org/10.1787/5k3v0hldmchl-en. [113]

Seidel, T. and R. Shavelson (2007), “Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results”, *Review of Educational Research*, Vol. 77/4, pp. 454-499, http://dx.doi.org/10.3102/0034654307310317. [7]

Skaalvik, E. and S. Skaalvik (2010), “Teacher self-efficacy and teacher burnout: A study of relations”, *Teaching and Teacher Education*, Vol. 26/4, pp. 1059-1069, http://dx.doi.org/10.1016/j.tate.2009.11.001. [54]

Teddlie, C. and D. Reynolds (eds.) (2000), *The International Handbook of School Effectiveness Research*, Routledge, New York, NY. [24]

Thapa, A. et al. (2013), “A review of school climate research”, *Review of Educational Research*, Vol. 83/3, pp. 357-385, http://dx.doi.org/10.3102/0034654313483907. [81]

Tibshirani, R. (1996), “Regression shrinkage and selection via the lasso”, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58/1, pp. 267-288, http://dx.doi.org/10.1111/j.2517-6161.1996.tb02080.x. [112]

Tschannen-Moran, M. and A. Hoy (2001), “Teacher efficacy: Capturing an elusive construct”, *Teaching and Teacher Education*, Vol. 17/7, pp. 783-805, https://doi.org/10.1016/S0742-051X(01)00036-1. [47]

Viac, C. and P. Fraser (2020), "Teachers' well-being: A framework for data collection and analysis", *OECD Education Working Papers*, No. 213, OECD Publishing, Paris, https://dx.doi.org/10.1787/c36fc9d3-en. [48]

Wang, M. and J. Degol (2016), "School climate: A review of the construct, measurement, and impact on student outcomes", *Educational Psychology Review*, Vol. 28/2, pp. 315-352, http://dx.doi.org/10.1007/s10648-015-9319-1. [125]

Wayne, A. and P. Youngs (2003), "Teacher characteristics and student achievement gains: A review", *Review of Educational Research*, Vol. 73/1, pp. 89-122, https://doi.org/10.3102/00346543073001089. [8]

Weiss, E. (1999), "Perceived workplace conditions and first-year teachers' morale, career choice commitment, and planned retention: A secondary analysis", *Teaching and Teacher Education*, Vol. 15/8, pp. 861-879, https://doi.org/10.1016/S0742-051X(99)00040-2. [82]

Wilder, S. (2014), "Effects of parental involvement on academic achievement: A meta-synthesis", *Educational Review*, Vol. 66/3, pp. 377-397, http://dx.doi.org/10.1080/00131911.2013.780009. [126]

Notes

[1] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[2] Based on the research literature, estimates of teacher effects (i.e. the difference between an average teacher and one at the 84th percentile), typically range from 0.1 to 0.2 standard deviations of student achievement (Jackson, Rockoff and Staiger, 2014[12]).

[3] PISA is a triennial survey of 15-year-old students that assesses the extent to which they have acquired key competencies essential for full participation in social and economic life (OECD, 2019[128]). PISA assessments do not just ascertain whether students can reproduce what they have learned, but they also examine how well students can extrapolate from what they have learned and apply their knowledge. Hence, while PISA focusses on students' competencies and on how well these competences are applied in different contexts, it may not reflect the curriculum for 15-year-old students. In each round of PISA, one subject, among reading, mathematics and science, is tested in detail. The main subject in 2018 was reading.

[4] Since Viet Nam does not have data on PISA test scores, it is not included in the analyses presented in Chapter 2.

[5] It has to be noted that the school level could also refer to system level depending on the school governance arrangements that are in place.

[6] Teachers' perceived disciplinary climate is included within the dimension of classroom practices, since it can be regarded as a partial measure for teachers' classroom management skills.

[7] TALIS covers the following types of assessment practices: administering own assessment, providing written feedback on student work in addition to marking, letting students evaluate their own progress and observing students when working on particular tasks and providing immediate feedback.

[8] Instruction characterised by cognitive-activation include asking students to solve problems that require them to think for an extended time, for which there is no immediately obvious solution or that can be solved in several different ways, or by using alternative routes and procedures. These instructional practices may also include having students reflect on and share with their peers the process through which they solved specific problems covered in class or assigned as homework as well as require students to apply what they have learned to new contexts (Echazarra et al., 2016[35]).

[9] TALIS asks teachers about their views ("strongly disagree"; "disagree"; "agree"; or "strongly agree") on innovation with four statements: "most teachers in this school strive to develop new ideas for teaching and learning"; "most teachers in this school are open to change"; "most teachers in this school search for new ways to solve problems"; "most teachers in this school provide practical support to each other for the application of new ideas".

[10] Based on the TALIS 2018 teacher questionnaire, reading teachers are defined as those who teach reading, writing and literature in the mother tongue, in the language of instruction, or in the tongue of the country (region) as a second language (for non-natives); language studies, public speaking. Science teachers are those who teach science, physics, physical science, chemistry, biology, human biology, environmental science, agriculture/horticulture/forestry.

[11] Based on PISA results, girls slightly outperformed boys in science, by only two score points, on average across OECD countries (OECD, 2019[64]).

[12] For instance, analysis based on TALIS data show that in the case of teachers' satisfaction with the profession only a small percentage (i.e. 4%) of the total variance comes from differences between schools (OECD, 2020[43]).

[13] Since the focus is on the average subject domain teacher within schools, teacher variables averaged only for subject domain teachers and the analyses by subject are based on samples restricted to schools with at least one subject domain teacher.

[14] In TALIS, the index of classroom disciplinary climate measures the extent to which teachers perceive disciplinary issues in the class. Higher values of the index of classroom disciplinary climate indicate a higher need in classroom discipline.

[15] In TALIS, the index of workload stress measures the extent to which workload, including lesson preparation, lessons to teach, marking, administrative work and extra duties due to absent teachers, is considered a source of stress. Higher values of the index of workload stress correspond to workload being considered a more important source of stress.

[16] In TALIS, the index of stakeholders' view of the value of the teaching profession measures the extent to which teachers feel that their views are valued by policy makers and the media in the country/region and that they can influence educational policy in the country/region.

[17] In TALIS, indicators related to classroom composition, including the average concentration of low academic achievers and of students from socio-economically disadvantaged backgrounds, are based on teachers' responses. For the analyses included in this report, these indicators on classroom composition are measured as the central values of percentage ranges: 0%, 5%, 20%, 45% or 80%.

[18] In TALIS, the index of stakeholder involvement measures principals' account on the extent to which the following statements apply to the school: parents/guardians support student achievement; parents/guardians are involved in school activities; and the school co-operates with the local community.

[19] In TALIS, the index of teacher-student relations measures how strongly teachers agree or disagree with the following statements about what happens in the school: teachers and students usually get on well with each other; most teachers believe that the students' well-being is important; most teachers are interested in what students have to say; and if a student needs extra assistance, the school provides it.

[20] The controls for the average classroom composition within the school (also referred to as classmates' characteristics) include the share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with migrant background and students who are refugees.

[21] The results of Argentina can be considered as a proxy for Ciudad Autónoma de Buenos Aires.

[22] Based on PISA 2018 results, for the Czech Republic and Turkey, the isolation index of high-achieving students in reading is between 0.35 and 0.4. This means that, in these countries, a student who scores in the top quarter of the distribution of PISA performance within a country has a more than one-in-two chance of attending the same school than students who are also high achievers, while this likelihood would have been only one in four if students had been uniformly distributed across schools (OECD, 2020[43]).

3 What do teachers and schools do that matters most for students' social and emotional development?

This chapter begins by reviewing a broad range of student attitudes, behaviours and aspirations towards school in an attempt to identify those likely to vary significantly between schools. It then focuses on four social-emotional outcomes for further analysis: students' perceptions of their classroom climate, teachers' enthusiasm for teaching, test performance, and students' educational expectations. Using a machine learning technique, lasso, and traditional regression analyses, it then aims to identify teacher and school dimensions that are the most significantly related to these four student social-emotional outcomes. For each of these outcomes, it attempts to identify cross-country patterns and differential teacher and school effects, as well as the mediating effects of classroom composition.

Highlights

- Schools and their teachers can actually make a difference for a wide range of students' social-emotional outcomes pertaining to student school behaviour, interest in school and well-being. More specifically, data from the nine countries and economies participating in the TALIS-PISA[1] link show that schools differ significantly in the way students perceive their classroom climate and their teachers' enthusiasm for teaching. There are also significant differences between schools in terms of students' perceptions of the difficulty of the PISA test and their expectations of completing at least a tertiary degree.
- A machine learning technique applied to TALIS-PISA link data retains four potential key predictors of these student social-emotional outcomes: teachers' classroom practices, classroom composition, school culture and school leadership. Variance decomposition analyses further indicate that four other dimensions explain an important part of differences in student social-emotional outcomes between schools: teachers' use of their working time, teachers' initial teacher education (both in terms of level and content), teachers' participation in professional development activities of a certain type and content, and teachers' well-being at work and their job satisfaction.
- On average across TALIS-PISA link countries and economies, students whose classmates include a substantial number of low academic achievers are less likely to expect to complete at least a tertiary education degree. This holds true while controlling for student's own socio-demographic background.
- Students who are enrolled in schools with larger classes tend to have higher educational expectations and to find the PISA test easier. This holds true while accounting for the academic and social-behavioural characteristics of the student's classmates. This could suggest that the larger the peer group is, the richer and more favourable the social-emotional development of the student. Yet, the presence of a selection bias cannot be ruled out. It is possible that schools implement specific grouping strategies.
- The way teachers use and allocate their working time across various professional activities contributes to the way students perceive their learning environment and shapes their educational expectations, at least in several participating countries. Specifically, the more teachers report spending time on marking and correcting student work and engaging in extracurricular activities, the better it is for students' social-emotional development.
- The more teachers there are in a school that hold a master's or a doctorate degree, the more likely students expect to complete a tertiary degree on average across TALIS-PISA link countries and economies. This might be because these teachers are better informed about the existing tracks for completing a tertiary degree, or because they represent higher educational aspirations for their students. However, highly-educated teachers might also be more likely to be employed in academically ambitious schools.
- On average across TALIS-PISA link countries and economies, students tend to find their teachers more interested in their teaching when those teachers report lower levels of work-related stress on average in their school.
- The findings highlighted above are established on average across TALIS-PISA link countries and for at least a couple of education systems. They cannot be interpreted as causal but, rather, correlational as TALIS and PISA measure teacher, principal, student and school characteristics in many countries at a single point of time. In addition, it is impossible to determine the sense of the relation, i.e. which factor drives changes in the other factor.

Introduction

What students experience and feel at school matters. It not only matters for student achievement but also for their well-being at school and beyond. This is primarily a source of concern for parents but also for teachers. The PISA 2018 report on what school life means for students notes that more than students' success at school it is their social and emotional development that is at the centre of parents' attention: "*When parents around the world are asked what they want for their children, some mention 'achievement' or 'success', but most reply 'happiness', 'confidence', 'friends', 'health', 'satisfaction', 'freedom from bullying' and the like."* (OECD, 2019, p. 40[1]). Turning to teachers, the TALIS 2018 results indicate that more than 90% of teachers in the OECD became teachers because teaching allowed them to influence the development of children and young people (OECD, 2019[2]). Students' school life matters for key student outcomes – students' academic success as well as social and emotional development – as these effects are mutually reinforcing and persist for years (Hoy, Hannum and Tschannen-Moran, 1998[3]).

Chapter 2 of this report has examined how teachers and schools contribute to students' performance in the PISA test following traditional conceptualisation of student academic achievement and empirical studies on the education production function (Hanushek and Rivkin, 2010[4]; Todd and Wolpin, 2007[5]). However, a substantial body of evidence indicates that student learning is multidimensional. Many factors beyond their core knowledge and skills are important contributors to both short- and long-term success and well-being. For example, psychologists find that personality influences the quality of one's thinking (Barratt, 1995[6]) and that grit and self-control influence how much a child learns in school (Duckworth et al., 2007[7]; Duckworth, Quinn and Tsukayama, 2012[8]). Longitudinal studies showed that childhood self-control, emotional stability, persistence, and motivation have long-term effects on health and labour market outcomes in adulthood (Borghans et al., 2008[9]; Chetty et al., 2011[10]; Moffitt et al., 2011[11]). Some studies even found that these sorts of attitudes and behaviours are stronger predictors of long-term outcomes like college attendance, earnings, home ownership and retirement savings than test scores (Chetty et al., 2011[10]).

Exactly like learning, teaching and schooling can also be characterised as multidimensional (Blazar and Kraft, 2017[12]). High-quality teachers and schools are expected not only to raise student performance, as measured through PISA and other tests, but also to provide emotionally supportive environments that contribute to students' social and emotional development (Blazar and Kraft, 2017[12]; Pianta and Hamre, 2009[13]). In recent years, two research approaches have tested this theory using empirical evidence. The first research approach focused on estimating teachers' contribution to student outcomes, often referred to as "teacher effects" or "teacher value-added" (Chetty, Friedman and Rockoff, 2014[14]; Hanushek and Rivkin, 2010[4]). These studies found that, as with test scores, teachers vary considerably in their ability to impact students' social and emotional development including a variety of observed school behaviours (Gershenson, 2016[15]; Jackson, 2018[16]; Jennings and DiPrete, 2010[17]; Koedel, 2008[18]; Kraft, 2019[19]; Ladd and Sorensen, 2017[20]; Ruzek et al., 2015[21]). Moreover, these studies also suggest that teachers who make a difference for students' academic performances are not necessarily equally good at contributing to students' social and emotional development, hence putting emphasis on the multidimensionality of teaching and schooling. The second research approach focused on observations of classrooms as a means of identifying domains of teaching practice that matter for students' cognitive as well as social-emotional outcomes (Blazar et al., 2017[22]; Hafen et al., 2015[23]). Several domains, such as teachers' interactions with students, classroom organisation, and emphasis on critical thinking within specific content areas, are found to support students' development in areas beyond their core academic skills (Blazar and Kraft, 2017[12]).

A few studies combine both approaches. This is the case of the OECD *Global Teaching InSights* (formerly referred to as the TALIS Video Study) (OECD, 2020[24]). This study used a pre/post design aimed to measure the impact of the teaching practices observed through videos on student outcomes (OECD, 2020, p. 23[24]). The study was more conclusive regarding what teaching practice matters for student social-emotional outcomes than for student achievement. The report showed that teachers' practices such as classroom management and the social-emotional support of students were significant predictors of student personal interest and self-efficacy in mathematics in half of countries/economies, even after accounting for students' prior mathematics performance and other background characteristics. In contrast, the quality of instruction was significantly associated with student achievement in one country only (OECD, 2020, p. 16[24]).

Chapter 3 of this report contributes to this research field by addressing this main research question: what do teachers and schools do that matters most for students' social and emotional development? The chapter's contribution differs from other research studies of this field in multiple ways. It draws on the rich TALIS-PISA 2018 link dataset of 15-year-old students, their teachers and schools. While most studies of this research field have been conducted mainly on local and state-level datasets in the United States, this chapter presents results at an international scale with data from nine countries and economies from four continents – Australia, Ciudad Autónoma de Buenos Aires (hereafter "CABA" [Argentina]), Colombia, the Czech Republic, Denmark, Georgia, Malta, Turkey and Viet Nam. Unlike past studies that usually focused on kindergarten, pre-primary and primary education, this chapter focuses on 15-year-old students and teachers teaching this age group. In comparison to past studies that usually analysed just a few student outcomes, this chapter reviews a wealth of student-reported indicators on dimensions of social and emotional development that include perceptions of their school's climate and their own contributions to shaping it, their well-being and their interest in school. All these measures are important outcomes of interest to policy makers, researchers, teachers and parents – see, in particular, Volume III of *PISA 2018 Results* by OECD (2019[1]). The large set of students' social-emotional outcomes is complemented by teachers' and principals' reports, collected as part of the rich TALIS 2018 questionnaire. These cover not only teaching practices but a wide range of aspects concerning their training, professional practices, working conditions and work environment.

The TALIS-PISA link data also present important limitations. The link between teachers and students is established at the school and not at the class level. In other words, the data do not allow a teacher to be matched with her/his students, but, rather, a sample of teachers teaching 15-year-old students in a school to be matched with a sample of 15-year-old students of that same school. Information on teachers is therefore averaged at the school level to be analysed together with students' outcomes. Given that teachers of the same school differ significantly in terms of their characteristics and practices, linking data by averaging teachers' variables at the school level triggers a considerable loss of information.[2] Last but not least, the cross-sectional design of the survey prevents the measurement of any teachers' causal effects. Neither does it allow for short-term effects of teachers and schools on students' outcomes to be distinguished from long-term ones.

The first section of this chapter reviews a broad range of student social-emotional outcomes to identify those that significantly vary between schools. The second section of the chapter focuses on the four student social-emotional outcomes that present important school variations: student perceptions of the disciplinary climate and their teacher's enthusiasm in their language-of-instruction lessons, student appreciation of difficulty of the PISA test and student educational expectation. The goal is to identify which teacher and school dimensions are the most significantly related with these four outcomes. To this end, Chapter 3 builds heavily on the conceptual framework and statistical methods presented in Chapter 2. In particular, it uses the same supervised statistical learning method, lasso, to investigate the relationship between teacher and school factors, and student socio-emotional development. It finally attempts to identify cross-country patterns, and differential teacher and school effects depending on the outcome of interest as well as the mediating effects of classroom composition.

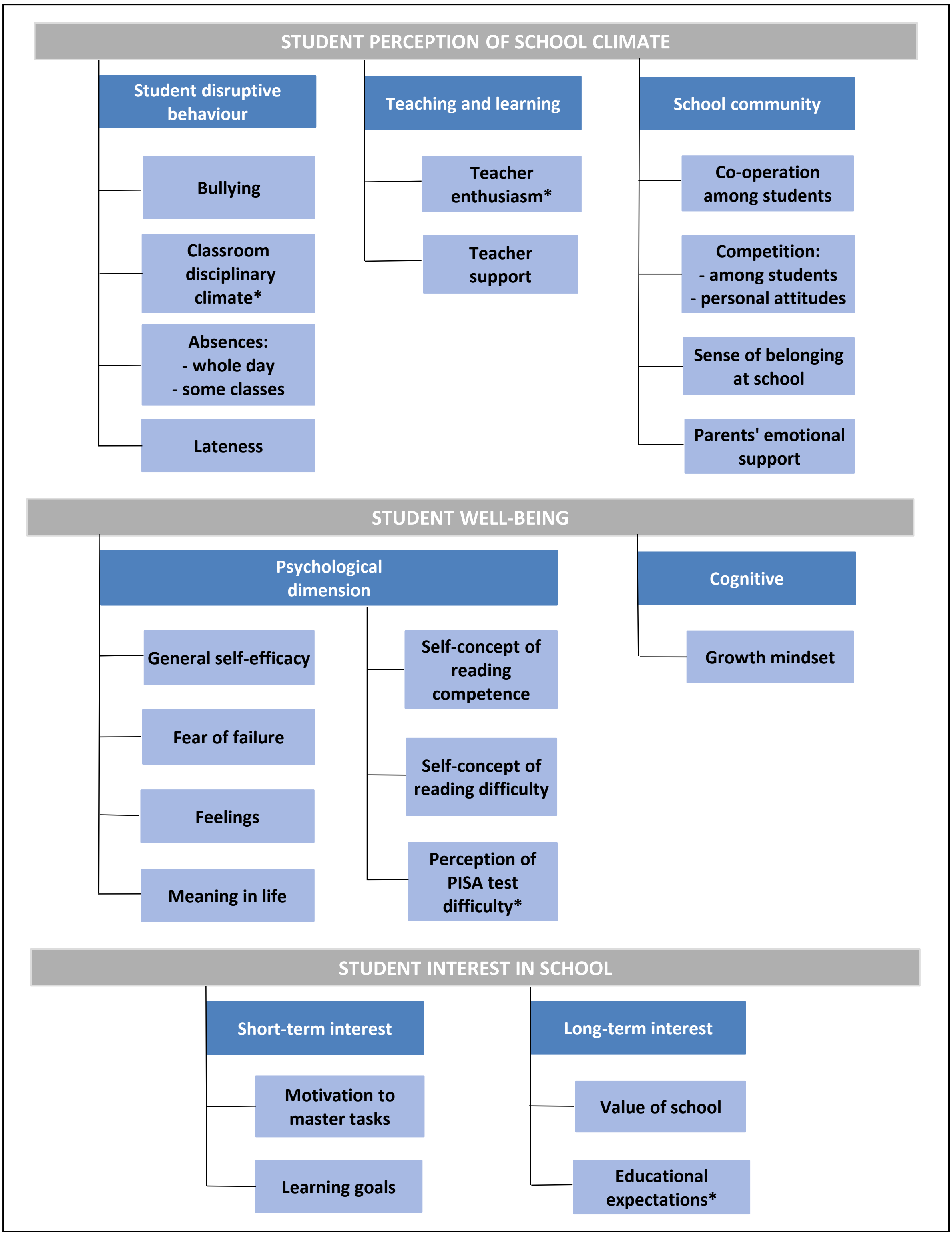
To what extent do schools matter for students' social and emotional development?

Several recent studies on teachers' "value-added" have documented the magnitude of teacher effects on a range of student outcomes other than test scores (Gershenson, 2016[15]; Jackson, 2018[16]; Jennings and DiPrete, 2010[17]; Koedel, 2008[18]; Kraft, 2019[19]; Ladd and Sorensen, 2017[20]; Ruzek et al., 2015[21]). These studies attempt to isolate the unique effect of teachers on social-emotional outcomes from factors outside of teachers' control (e.g. students' prior achievement, gender, socio-economic and cultural status, immigration background) and to limit any bias due to non-random sorting of teachers and students. These studies found small-to-important teacher effects on many kinds of students' social-emotional outcomes.

This chapter first reviews a broad range of student attitudes, behaviours and aspirations towards school to identify those that are more likely to be affected by what a school, its teachers and its students do. This chapter supplements the conceptual frameworks for school climate and students' well-being put forward in *PISA 2018 Results (Volume III): What School Life Means for Students' Lives?* (OECD, 2019[1]) with a conceptualisation of student interest in school – key to students' longer-term academic success (Hoy, Hannum and Tschannen-Moran, 1998[3]). Students' well-being, their interest in school, and perception of school climate and the part they play in it, can all be considered an integral part of students' social-emotional development. Indeed, the OECD defines social and emotional skills as: "…individual capacities that can be (a) manifested in consistent patterns of thoughts, feelings and behaviours, (b) developed through formal and informal learning experiences, and (c) important drivers of socioeconomic outcomes throughout the individual's life" (OECD, 2015, p. 35[25]). In this chapter, a forward-looking perspective is taken. It considers the outcomes of interest as indicators, at a given time, of an in-progress development of student social-emotional skills.[3]

The first part of this chapter maps social-emotional outcomes measured by PISA up against outcomes that studies on teachers' value-added have identified as being affected by teachers. It then estimates how much of the differences in each student outcome can be attributed to differences between schools. The differences observed between schools can result from different sources: actual differences between schools in their school body and peer-to-peer interactions, in the school's staff, its working conditions and practices as well as in the school's characteristics and policies. This first part aims at delimiting what can actually be analysed and concluded from the TALIS-PISA link data, i.e. through a school-level link of cross-sectional data. Indeed, only a few of the 24 student social-emotional outcomes are found to be subject to significant school differences and, even in such cases, school differences remain weak in comparison to differences observed within schools among students. This substantially limits identification of what teachers and schools do that matters for student social and emotional development, and should be borne in mind for the second part of this chapter. The second part strives to identify teacher and school dimensions that matter for the student outcomes retained for analysis.

Figure 3.1. Student social and emotional outcomes as measured in TALIS-PISA link 2018



Note: Social-emotional outcomes flagged with an asterisk (*) are those for which a significant share of their total variance (i.e. more than 10% of the total variance) lies between schools in all or almost all participating countries.
Source: Adapted from OECD (2019[1]), *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*, PISA, https://dx.doi.org/10.1787/acd78851-en, Figures III.1.1 and III.1.2.

Student perception of and part played in school climate

Among all aspects pertaining to school climate (see the upper part of Figure 3.1), student disruptive behaviour is the one aspect that was found to be subject to teacher effects by past research studies. Some studies did find teachers' effects on students' absences (Gershenson, 2016[15]; Jackson, 2012[26]; Ladd and Sorensen, 2017[20]) as well as on students' behaviour in class and suspensions (Blazar and Kraft, 2017[12]; Jackson, 2018[16]; Ladd and Sorensen, 2017[20]), while Jennings and DiPrete (2010[17]) also found an effect of kindergartens' staff on children's self-control. The fact that none of these studies found teacher effects on other indicators of school climate, such as a student's sense of belonging at school or propensity to co-operate or collaborate, does not necessarily result from the absence of such effects. On the contrary, this can result from a lack of such analysis or data available for these indicators.

The TALIS-PISA link data do not allow measurement of teacher value-added because of their cross-sectional design and school-level link. Yet, the TALIS-PISA link dataset allows identification of the extent to which schools (including what makes up a school – its teacher workforce and its students) differ for a varied range of student outcomes. Analyses conducted on the TALIS-PISA link data indicate that, out of the 12 student outcomes pertaining to school climate, two vary significantly across schools (they are flagged with an asterisk * in Figure 3.1), meaning that a significant share of the total variance in two of these outcomes (i.e. more than 10% of the total variance) lies between schools in all or almost all participating countries (Table 3.1). These are students' perceptions of classroom disciplinary climate[4] and teachers' enthusiasm for teaching.[5]

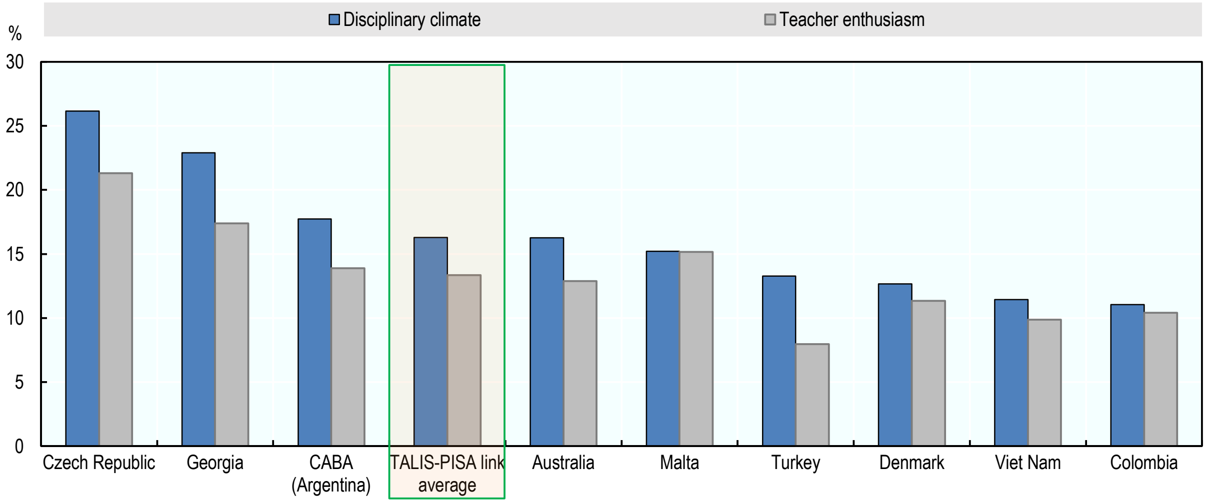
Schools significantly differ in the way students report how they behave in class. The between-school variance of student perception of the classroom disciplinary climate ranges from 11% in Colombia and Viet Nam to 23% in Georgia and even 26% in the Czech Republic (Figure 3.2). This finding resonates with those established by Blazar and Kraft (2017[12]). They found that upper-elementary teachers have large effects on self-reported measures of students' behaviour in class in four districts on the east coast of the United States. They also estimated that the variation in teacher effects on students' behaviour in class is of similar magnitude to the variation in teacher effects on mathematics scores. This might suggest that the school differences found in classroom climate for TALIS-PISA link countries and economies result from cumulative teacher effects on student behaviour in class. At the same time, differences in classroom climate between schools relate largely to differences in performances. In the case of the Czech Republic, the large differences observed in disciplinary climate between schools partly result from the tracking structure of the Czech education system with schools in the general tracks enjoying a more favourable school climate than vocational schools.

Schools also significantly differ in the way students perceive their teachers' enthusiasm for teaching. The between-school variance of this index ranges from 8% in Turkey and 10% in Colombia and Viet Nam to 21% in the Czech Republic (Figure 3.2).

However, the share of between-school variance is relatively low for the other ten social-emotional outcomes, including indicators on the concept of school community (e.g. student co-operation and competition and student sense of belonging at school) (Table 3.1). Also, one could have expected to find significant differences between schools with regard to student truancy, at least for two reasons: some schools could have a higher concentration of students with deviant behaviours than others because of school and residential segregations. Second, important teacher value-added was found in past US studies on student absences, suggesting important school differences in student absences if teachers' effects add up within schools (Gershenson, 2016[15]; Jackson, 2012[26]; Ladd and Sorensen, 2017[20]). Unfortunately, it is impossible to disentangle the reasons for this unexpected finding.

Figure 3.2. Schools differ in student perception of classroom climate and teacher enthusiasm

Percentage of total variance in the indices of classroom disciplinary climate and teacher enthusiasm, explained at the school level



Note: The figure shows the R^2 of the linear regression of student perception of classroom climate (respectively of teacher enthusiasm) on school fixed effects.
Countries and economies are ranked in descending order of the percentage of total variance in the index of classroom disciplinary climate, explained at the school level.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 3.2 and 3.4.

StatLink https://doi.org/10.1787/888934223859

Student well-being

In the framework presented in the third volume of *PISA 2018 Results* on students' school life (OECD, 2019[1]), student well-being is composed of two sub-dimensions,[6] each of which includes several indicators that are relevant to this chapter (see the middle part of Figure 3.1):

- psychological dimension: general self-efficacy, self-concept of reading competence and difficulty, perception of difficulty of the PISA test*, fear of failure, feelings and meaning in life
- cognitive dimension: growth mindset.

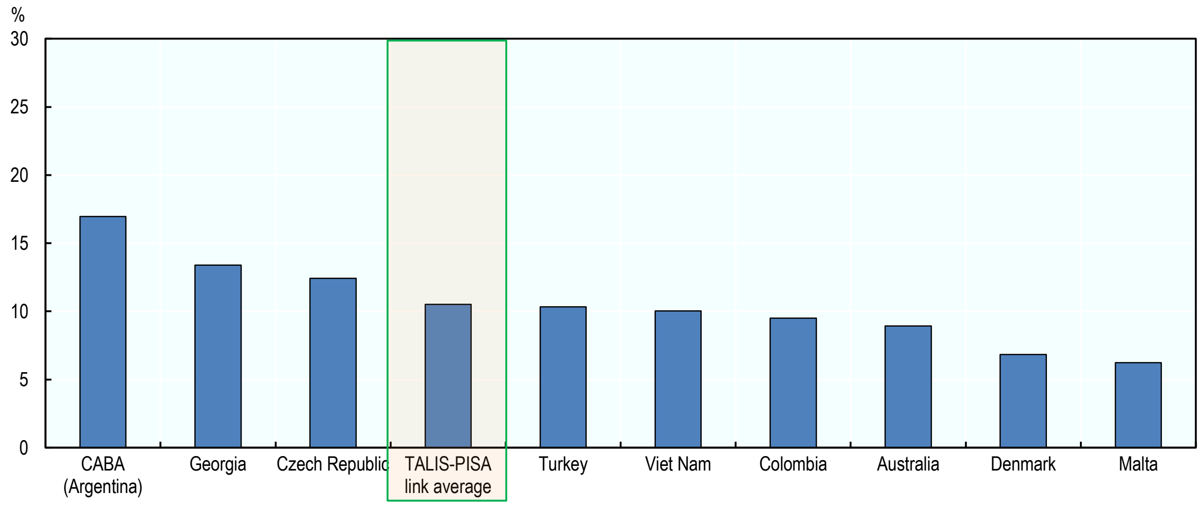
A few past studies examined whether teachers can have an impact on student well-being measures. Using data on upper-primary teachers collected by the National Centre for Teacher Effectiveness, Blazar and Kraft (2017[12]) estimated that the variation in teacher effects on students' self-efficacy in mathematics is of similar magnitude to the variation in teacher effects on mathematics test scores. They found that the variation of teacher effects on students' happiness in class is even larger. Using data on upper-primary teachers collected by the Measures of Effective Teaching (MET) Project, Kraft (2019[19]) found significant teacher effects on student growth mindset. Overall, PISA measures of student well-being map well with student outcomes identified by research as influenced by teachers, and even extend their scope.

As mentioned before, the TALIS-PISA link data do not allow measurement of teacher value-added. Yet, it is possible to identify the extent to which schools differ in their students' well-being. Analyses indicate that of the eight student outcomes related to well-being only students' perception of the difficulty of the PISA test[7] significantly varies between schools (flagged with a * above and in Figure 3.1) (Table 3.1). This means that a significant share of the total variance (i.e. more than 10% of the total variance) lies between schools in all or almost all participating countries.

School differences in the way students perceive PISA test difficulty are the weakest in Malta (6%) and Denmark (7%), culminate at 17% of the total student differences in CABA (Argentina) and amount to around 10% in all remaining countries and economies (Figure 3.3). One may wonder why this student outcome is more subject to differences between schools than others such as student general self-efficacy or growth mindset. A possible explanation for this is that students from the same classes and schools tend to perform similarly – the between-school variance in student performance (around 30%, see Chapter 2) is actually higher than that in student perception of test difficulty. Moreover, students attending the same classes and schools can sit the same school assessments or tend to have similar testing experiences and therefore similar appreciation of the PISA test difficulty. Other indicators of well-being, such as students' feelings or sense of meaning in life, are not directly related to school, and are general measures of individual well-being. Therefore, they might be less subject to between-school variations. In addition, in comparison to student general self-efficacy or self-concept in reading, the question on the PISA test difficulty asks about a more concrete situation that just happened, and results in more precise estimates of student outcomes at student and school levels, thereby impacting between-school variance.

Figure 3.3. Schools differ in student perception of difficulty of the PISA test

Percentage of total variance in the index of student perception of difficulty of the PISA test, explained at the school level



Note: The figure shows the R^2 of the linear regression of student perception of difficulty of the PISA test on school fixed effects.
Countries and economies are ranked in descending order of the percentage of total variance in the index of student perception of difficulty of the PISA, explained at the school level.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.6.

StatLink https://doi.org/10.1787/888934223878

Student interest in school

In addition to the two sets of student outcomes presented above – perception of and participation in school climate and student well-being – this chapter examines student interest in school, which is key to longer-term educational success. This includes the following indicators:

- students' motivation to master tasks
- students' learning goals
- student perception of the value of school and
- students' educational expectations*.

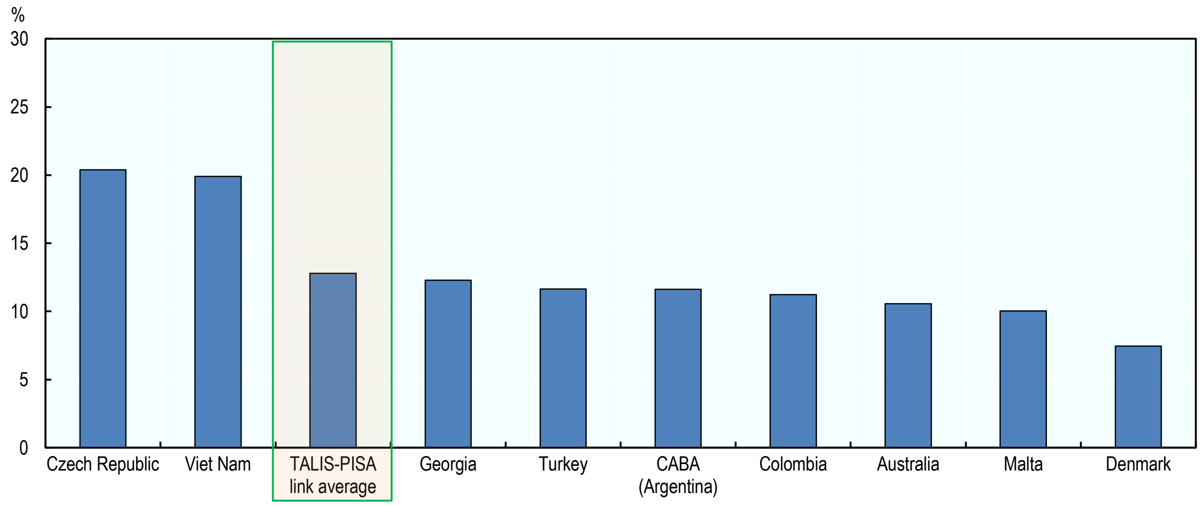
It is possible to map some PISA indicators pertaining to student interest in school to student outcomes that were found to be impacted by teachers from studies on teachers' value-added. Jennings and DiPrete (2010[17]) estimated the role that teachers play in developing kindergarten and first-grade students' social and behavioural outcomes, including the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility and organisation. They found within-school teacher effects on social and behavioural outcomes that were even larger than effects on students' academic achievement. In a study of middle-school mathematics teachers, Ruzek et al. (2015[21]) found small but meaningful teacher effects on students' motivational orientation (mastery and performance achievement goal) among seventh graders. Yet, the analysis suggests less variation in teachers' contributions to students' achievement goals than mathematics achievement. A few studies found teacher effects on students' educational aspirations, especially on female students' aspirations to pursue studies in science, technology, engineering and mathematics (STEM) (Breda et al., 2020[27]; Carrell, Page and West, 2010[28]). A couple of studies also analysed teacher effects on long-term educational outcomes such as college attendance. Using data from primary teachers, Chetty et al. (2014[14]) found that students assigned to high value-adding teachers are more likely to attend college. Jackson (2018[16]) found that teachers have causal effects on skills not measured by testing, but reflected in absences, suspensions, grades, and on-time grade progression. Teacher effects on these non-test-score outcomes in the 9th grade predict longer-run effects on high-school completion and proxies for college-going above and beyond their effects on test scores. Effects on non-test-score outcomes were found to be particularly important for English teachers.

Analyses conducted on the TALIS-PISA link data indicate that of the four student outcomes pertaining to student interest in school, one of them – students' educational aspirations (flagged with a * above and in the bottom part of Figure 3.1) – differs significantly between schools (Table 3.1). More specifically, PISA measures education expectations by asking students which educational level they expect to complete. Their responses were used to create a dummy variable that equals 1 if the student expects to complete at least a tertiary degree and 0 otherwise.[8]

Schools differ more in students' educational expectations in some countries than in others. The between-school variance[9] for this outcome amounts to 7% in Denmark, culminates at 20% in the Czech Republic and Viet Nam, and ranges between 10% and 12% for all remaining countries and economies (Figure 3.4). The large differences observed between schools in the case of the Czech Republic must partly result from the structure of the education system, which enrols 15-year-old students in different tracks that lead students to complete studies differing in length and type of content. Educational expectations are more subject to differences between schools than other measures of students' interest in school[10] (Table 3.1). An explanation for this might be that educational aspirations are more stable in time as they refer more to student long-term interest in school than the determination and motivation students report in pursuing their day-to-day school work. It might also be the case that the question about educational expectations is more concrete than the others as they refer to extant educational degrees rather than values. For both reasons, the indicator of student educational aspirations might yield more precise estimates of student outcomes at the student level and therefore at the school level, which matters for the estimation of the between-school variance.

Figure 3.4. Schools differ in student educational expectations

Proxy for the percentage of total variance in student expectation of completing at least a tertiary degree, explained at the school level



Note: In the case of binary variables analysed with the help of logistic regressions, an equivalent statistic to R^2 does not exist. The goodness-of-fit of logistic regression models is evaluated by pseudo R^2. Pseudo R^2 is on a similar scale to R^2 (ranging from 0 to 1) with higher values indicating better model fit. It is used as a proxy for the percentage of the variance in the dependent variable that the independent variables explain collectively.
Countries and economies are ranked in descending order of the percentage of pseudo R² of the logistic regression of student expectation of completing at least a tertiary degree on school fixed effects.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.7.

StatLink https://doi.org/10.1787/888934223897

Summary and discussion

Interestingly, the four social and emotional outcomes that present significant differences between schools belong to all of the three general concepts presented above – school climate, student well-being and student interest in school. This shows that teachers and schools can actually matter for a wide range of students' outcomes. At the same time, only four of the 24 indicators analysed are found to vary significantly across schools. And, even for these indicators, school differences are relatively weak, particularly in comparison with differences in achievement between schools (see Chapter 2). Yet, though school systems and schools may effectively manage students' social-emotional development, this should not be overstated.

First, research conducted in the United States suggests that teacher effects on student achievement are similar in magnitude if not smaller than teachers effects on certain student attitudes and behaviours (Blazar and Kraft, 2017[12]). Little is known about the way teacher effects aggregate at the school level, i.e. whether they tend to cumulate or cancel each other out within schools, and whether this varies depending on the outcome of interest. A plausible reason why schools appear to differ more with respect to student achievement than student social-emotional outcomes has to do with the fact that student achievement is more precisely measured than student social-emotional outcomes in PISA. While student performances in each domain are evaluated based on at least a half an hour of test-taking, indicators of student social-emotional development are measured based on one question each.[11]

In the same way, it can be argued that the four social-emotional outcomes that show the largest differences between schools do so because they are more precisely measured. Two of these four outcomes refer to what happens specifically in language-of-instruction lessons. These are students' perceptions of the classroom disciplinary climate and the teacher's enthusiasm in language-of-instruction lessons. This might be because students report on aspects pertaining to a collective environment – the classroom and its teacher – and not on their own personal behaviour, beliefs and values. It is thus likely that students' reports on this shared reality are more similar within schools as students from the same classroom and/or taught by the same teacher can be surveyed by PISA. The other two outcomes specifically refer to more objective evaluation of student self-concept and interest in school: student perception of the PISA test difficulty and student expectation of completing at least a tertiary education degree.

In addition, because these four outcomes are more concrete, they are also less likely to suffer from reference bias (Van de gaer et al., 2012[29]) than the other 20 outcomes under scrutiny. When students rate themselves on statements such as: "I am a good reader"; "I am able to understand difficult texts": or "I normally feel happy"; "I normally feel proud", they must interpret the definitions of reading proficiency or happiness, which likely involves comparing themselves with other people, in particular with students from their class or their school. Thus, it is likely that students from different schools have different standards or reference points when answering less concrete questions. In such cases, the differences between schools in the social-emotional outcomes that suffer (the most) from reference bias are underestimated.

Finally, there are some consistent cross-country patterns that emerge from the analyses across these four outcomes. Irrespective of the social-emotional outcome of interest, the Czech Republic and Georgia are consistently among the three countries with the largest school differences. At the other end of the spectrum, Colombia and Denmark are consistently among the four countries showing the smallest school differences (yet still at around 10% of the total variance or above). These cross-country differences may result from two factors that can reinforce or partially mitigate each other: higher segregation of students by prior social-emotional skills and higher variations in school effects on student social-emotional outcomes in the Czech Republic and Georgia in comparison to other countries (especially in Colombia and Denmark). Unfortunately, the TALIS-PISA link data do not allow disentangling these effects. The structure of the education system, in particular the number of education programmes available to 15-year-olds and students' age at first selection, might explain part of the differences resulting from school segregation. For example, the large differences observed in the Czech Republic might be explained by the fact that 15-year-old students can be enrolled in five different education programmes and they are first selected into these tracks at the age of 11. On the other hand, in Denmark, where small differences are observed between schools, the age at first selection is 16 and all 15-year-old students are enrolled in the same education programme – see OECD (2020[30]), Figure V.3.2.

The following analyses aim to identify teacher and school factors that can explain the differences found between schools in student social-emotional development. These might be considered even more relevant for countries like the Czech Republic and Georgia where those differences are particularly large in comparison with other participating countries and economies.

To what extent do specific teacher and school factors matter for students' social and emotional development?

The remainder of the chapter focuses on the four student social-emotional outcomes that present important school variations: student perception of their language-of-instruction lessons (disciplinary climate and teacher enthusiasm), student perception of PISA test difficulty, and student educational expectation. It now aims at identifying which teacher and school dimensions actually matter for these four outcomes. This section of the chapter builds heavily on concepts and tools presented in Chapter 2. While Chapter 2 focused on student performance, this chapter focuses on student social-emotional outcomes that significantly vary across schools.

Teacher and school factors selected by lasso, a supervised statistical learning method

As in Chapter 2, this section uses a supervised statistical learning method, lasso, to investigate the relationship between teacher and school factors, and student social-emotional outcomes (see Box 2.1 for more information). It also borrows from the conceptual framework presented in Chapter 2 and uses the same list of potential predictors (almost 150 predictors across 18 teacher and school dimensions) to explain the four student outcomes under scrutiny.

Figure 3.5 provides an overview of teacher and school dimensions that are found to be significantly related to the four student social-emotional outcomes by focusing on two groups of teachers: all the teachers of the school irrespective of the subject they teach; and teachers teaching reading, writing or literature (henceforth "reading teachers"). As two of the four student outcomes under scrutiny refer to what happens in language-of-instruction lessons, it makes sense to replicate analyses conducted for all teachers on a sample restricted to reading teachers only. Also, as the major domain assessed in PISA 2018 is reading, it is interesting to replicate analyses of student perception of the PISA test difficulty while focusing on reading teachers only. Analyses of students' educational expectations are not restricted to the sample of reading teachers as this particular group of teachers is not expected to have a specific role in shaping students' educational expectations. That said, the remainder of the chapter focuses on the findings observed for the full sample including all teachers of the school as this larger sample of observations (all schools are included regardless of whether at least one reading teacher was surveyed or not) tends to trigger more significant results.

Lasso results suggest that, considering the overall population of 15-year-old students, teachers and principals surveyed within the TALIS-PISA link, none of the 18 teacher and school dimensions examined matter consistently for all four student social-emotional outcomes under scrutiny. This may point to the multidimensionality of teacher and school effects on student social-emotional outcomes since different dimensions of teachers' work and schools' practices seem to matter for different student outcomes. Yet the lack of significance in the findings might also result from a lack of statistical power or a loose link between the teachers and their students.

A few indicators are identified consistently by lasso as having a significant association with two to three student social-emotional outcomes (Figure 3.5). These indicators belong to four dimensions: teachers' classroom practices, classroom characteristics, school culture and school leadership.

Teachers' classroom practices

Results from lasso indicate that teachers' classroom practices are important for two of the four student social-emotional outcomes: student perception of classroom disciplinary climate and teacher enthusiasm in language-of-instruction lessons. This finding is aligned with results from past studies that used observation instruments to examine teachers' roles in supporting students' social-emotional development. One instrument in particular, the Classroom Assessment Scoring System (CLASS) collects data on aspects of teachers' instruction: teachers' social and emotional interactions with students, their ability to organise and manage the classroom environment, and their instructional supports in the delivery of content (Hafen et al., 2015[23]). A number of studies from developers of the CLASS instrument and their colleagues have described relationships between these dimensions and closely related student attitudes and behaviours. In particular, they found that teachers' classroom organisation predicts students' engagement and behaviour in class – see Blazar and Kraft (2017[12]) for a short review. The OECD *Global Teaching Insights* report showed that teachers' practices such as classroom management and their social-emotional support of students are significant predictors of student personal interest and self-efficacy in mathematics in half of the countries/economies participating in the study, even after accounting for students' prior mathematics performance and other background characteristics (OECD, 2020[24]).

Figure 3.5. Teacher and school factors that matter for student social-emotional outcomes, based on lasso

Dimensions and variables selected by lasso regressions based on responses of the overall population of 15-year-old students, teachers and principals surveyed within the TALIS-PISA link, by social-emotional outcome and by group of teachers

Dimensions	Variables selected	Students' perceptions of classroom climate		Students' perceptions of the PISA test difficulty		Students' perceptions of teachers' enthusiasm for teaching		Educational expectations
		All teachers	Reading teachers	All teachers	Reading teachers	All teachers	Reading teachers	All teachers
		(sign of standardised coefficients: +/-)						
Teacher dimensions with a direct effect on student achievement								
Classroom practices*	Use of class time (%): Keeping order in the classroom		-					
	Teachers' autonomy over planning and teaching	+	+					
	Teachers' perceived disciplinary problems *(higher values indicate higher need for classroom discipline)*	-	-			-	-	
Teacher dimensions with an indirect effect on student achievement								
Teacher characteristics	Years of experience as a teacher					-		
Motivation to join the profession	*No variables selected*							
Initial education and training	Highest level of formal education completed: Below ISCED 5						-	
	Content: Pedagogy of some or all subject(s) I teach					+		
Well-being and job satisfaction	Workplace well-being and stress					-		
	Workload stress *(higher values indicate workload being considered a more important source of stress)*							+
	Job satisfaction with work environment				-			
	Teachers' views of the way different stakeholders value the profession							-
Self-efficacy	Self-efficacy in classroom management						+	
	Self-efficacy in student engagement		+					
Working hours	*No variables selected*							
School dimensions with a direct effect on student achievement								
Classroom characteristics *(classmates' characteristics and class size)**	Classroom composition (%): Students whose first language is different from the language(s) of instruction	+						
	Classroom composition (%): Low academic achievers	-	-	+	+			-
	Classroom composition (%): Students with behavioural problems	-		+	+			
	Classroom composition (%): Students from socio-economically disadvantaged homes				+			
	Classroom composition (%): Academically gifted students			-				
School culture*	Stakeholder (i.e. parents and local community) involvement in school	+			-			
	Teacher-student relations	+				+		
School dimensions with an indirect effect on student achievement								
Induction	Took part in informal induction activities during first employment		-					
Mentoring	Having an assigned mentor					+		
Professional development	Content: Teaching students with special needs							-
Feedback	*No variables selected*							
Collaboration	*No variables selected*							
School innovativeness	*No variables selected*							
Employment status	Part-time (less than 50% of full time hours)		-					
Formal appraisal	*No variables selected*							
School leadership*	Principals' leadership activities: Collaborated with teachers to solve classroom discipline problems	-	-			-	-	

Notes: Country fixed effects and student characteristics, such as gender, immigrant background and index of economic, social and cultural status (ESCS), are always included among the variables selected by lasso.
Teacher variables are averaged at the school level for two different groups of teachers: all teachers in the school and all reading teachers in the school. Thus, the analyses focusing on reading teachers are based on samples restricted to schools with at least one reading teacher participating in TALIS.
Dimensions and variables that are identified consistently by lasso as having a significant association with two to three student social-emotional outcomes are highlighted in blue. Dimensions that are not selected for any of the social-emotional outcomes are highlighted in light grey. Since lasso is applied as a model selection technique it does not provide the standard errors required for performing statistical inference. The interpretation of the estimated standardised coefficients is conditional on the selected model and cannot be interpreted as causal. Moreover, in the presence of correlated explanatory variables, the signs of the coefficient estimates can swing based on which other independent variables are in the model.
For additional information on the full list of potential variables included in the lasso regressions and more generally about lasso, see Annex B.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.8.

Within this dimension of teachers' classroom practices, what lasso regressions particularly retain is that teachers' own perceptions of the disciplinary climate of the classroom is correlated with students' perceptions of both the classroom disciplinary climate and teachers' enthusiasm to teach. The more teachers report disciplinary problems in their classroom, the less students speak positively about their classroom disciplinary climate (hence the negative relationship show in Figure 3.5). This suggests that the different stakeholders are aligned in their perceptions of their shared environment but, also, potentially, that there is a virtuous/vicious spiral where perceptions and attitudes of stakeholders mutually reinforce each other. Lasso results focusing on reading teachers also show that the more time reading teachers spend on keeping order in their classroom, the more negative the disciplinary climate in the classrooms is as perceived by students. Lasso results further suggest that teachers who have more autonomy over lesson planning and teaching strategies tend to teach in classrooms with more disciplined students (Figure 3.5).

Classmates' characteristics and class size

Results from lasso regressions indicate that classroom characteristics are significantly related to three of the four outcomes of interest: the classroom disciplinary climate, students' perceptions of difficulty of the PISA test and students' educational aspirations (Figure 3.5). Since the lasso regressions control for students' own characteristics (gender, socio-economic and immigration background), these findings actually point to important "peer effects" on student social-emotional outcomes.

Low academic achievement and behavioural problems are the academic and socio-behavioural characteristics among classmates that matter the most.[12] The higher these shares are, the less disciplined classrooms are and the more students perceive the PISA test as difficult. These findings somewhat echo education research on peer effects (Sacerdote, 2011[31]) although most studies focus on peer effects on student achievement rather than on social-emotional outcomes. For example, Carrell and Hoekstra (2010[32]) showed that the classroom presence of children, in particular boys, exposed to domestic violence not only lowers maths and reading test scores of other children but also raises classroom discipline problems. They found that adding an additional troubled boy to a classroom raises the probability that another boy will commit a disciplinary infraction by 17% (and lowers test scores by two percentile points).

The classroom presence of low academic achievers is also related to students' own educational expectations. The higher the share of low academic achievers, the fewer the number of students who expect to complete at least a tertiary degree. This is somehow consistent with past research studies finding important peer effects on college completion. In particular, Gaviria and Raphael (2001[33]) found strong peer effects among tenth-graders in the likelihood of dropping out of high school as well as other social behaviour such as drug use, alcohol drinking, cigarette smoking and church going. In other words, students are more likely to engage in these activities the more their classmates do.

Identifying the precise channel(s) through which a given peer effect operates is an impossible task and is too demanding regarding the analytical potential of the TALIS-PISA link data. If a student's classmates have lower ability and are more disruptive in the classroom, the student herself/himself is more prone to create disruptions in the classroom and teachers may struggle more to keep order in the classroom. Also, if the student's classmates are low academic achievers, teachers likely teach at a slower pace and have lower expectations of their students' educational orientation. This, in turn, leads the student to have poorer educational expectations for herself/himself.

School culture and school leadership

Results from lasso regressions indicate that teacher-student relations, and the involvement of parents and the local community in school (Figure 3.5) are the elements of school culture that matter particularly for student social-emotional development. They are particularly associated with ensuring a disciplined classroom climate. Positive teacher-student relations go hand-in-hand with student perception of reading teachers' enthusiasm for teaching. Students also tend to find the PISA test less difficult when parents and the local community are involved in school (this only holds true when the analyses focus on reading teachers only). Parents' involvement in school may positively affect students' self-concept and perceptions of their class and teacher through several channels: greater collaboration between teachers, school staff and parents resulting in more tailored support for student social-emotional development; greater student perception of the value of school; and students' greater sense of co-operation and belonging to school.

Lasso regressions retain school leaders' involvement in solving classroom discipline problems as an important covariate of student perception of the disciplinary climate and teacher enthusiasm for teaching compared to other aspects of school leadership. This suggests consistency between students' and the principal's perceptions of the overall class and school climate. The more frequently principals report collaborating with teachers to solve discipline problems, the more students report disciplinary issues in their classrooms, and the less enthusiastic teachers are about teaching, according to students. Repeated disciplinary issues may be conducive to discouragement among concerned teachers.

Differences in student social-emotional outcomes explained by teacher and school factors

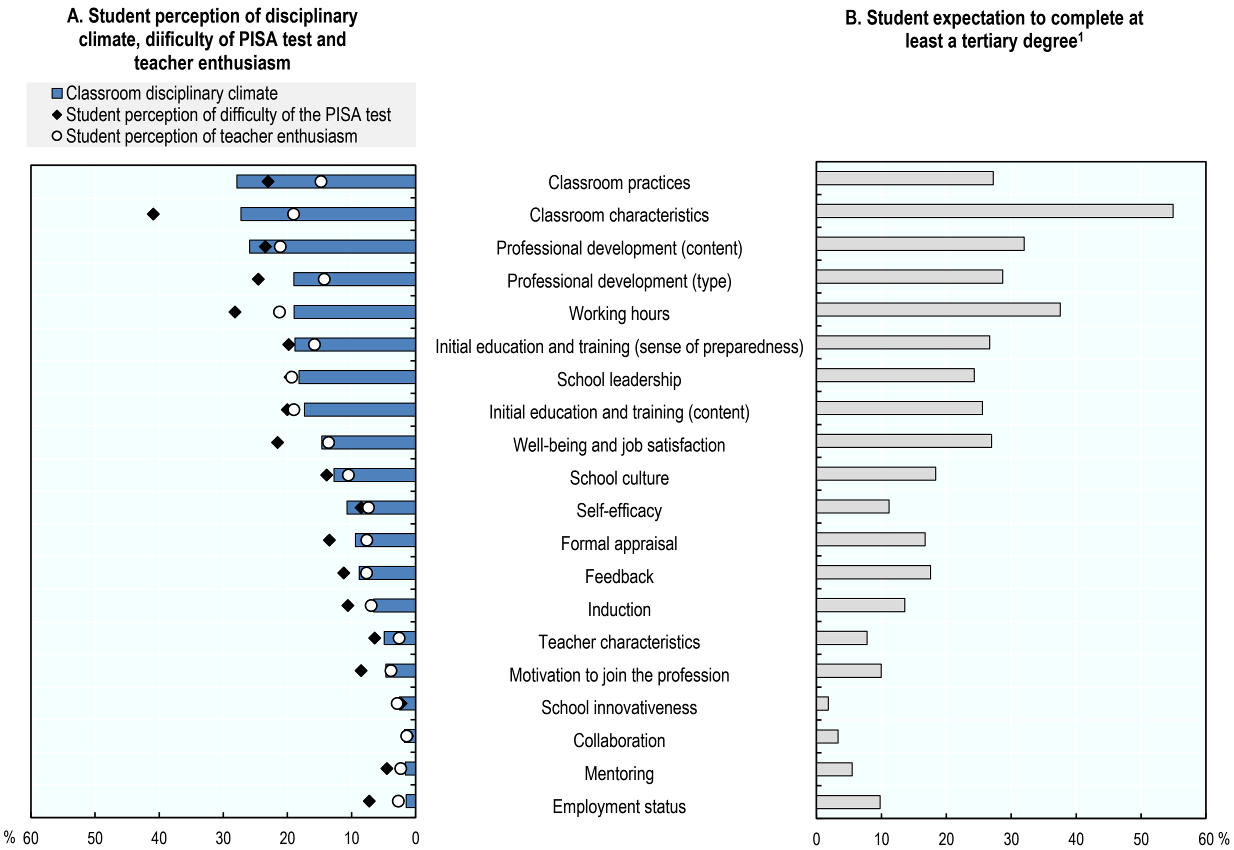
This chapter aims to identify teacher and school dimensions that matter the most for student social-emotional outcomes. Analysing results from lasso regression is one way to do this (see previous section). Another complementary approach consists of retaining those dimensions that explain the highest shares of variance in student social-emotional outcomes. Obviously, as the factors of interest included in the analyses are introduced at the school level, one can only explain the share of variance in students' social-emotional outcomes that lies at the school level. School average differences in student social-emotional outcomes across the nine TALIS-PISA link countries and economies represent about 10 to 20% of the total variance in student social-emotional outcomes (Figures 3.2, 3.3 and 3.4). This means that any of the teacher and school dimensions included in the analyses can only explain up to 10-20% of the total variance in student social-emotional outcomes.

The four dimensions previously identified as important for student social-emotional outcomes by lasso regressions also explain a substantial share of the differences between schools in student average social-emotional outcomes (Figure 3.6). Depending on the outcome of interest, the dimension of teachers' classroom practices explains between 15% (student perception of teacher enthusiasm) and 28% (classroom disciplinary climate) of school differences on average across participating countries and economies. Yet, classroom composition tends to explain a larger share of these differences – up to 55% of the total differences between schools in the share of students expecting to complete at least a tertiary degree. This corroborates lasso findings and past education research indicating important peer effects on students' social-emotional development. School leadership also explains around 20% of the between-school variance across all four student social-emotional outcomes. Yet, school culture is the one dimension among the four that captures the smallest share of the between-school differences in student social-emotional outcomes, though still above 10%.

Beyond these four dimensions flagged by lasso, there are four other dimensions that are found to explain an important part of the differences between schools in student social-emotional outcomes: teachers' use of working time; teachers' initial teacher education in terms of level and content; teachers' participation in professional development activities of certain type and content; and teacher well-being and job satisfaction (Figure 3.6). All of these dimensions capture a share of between-school variance that is around or above 20%. This warrants a detailed look at the indicators from these additional dimensions. In the next section, student outcomes are regressed on indicators of each teacher and school dimension taken separately.

Figure 3.6. Average school differences in student social-emotional outcomes explained by teacher and school factors

Percentage of between-school variance in student outcomes, by dimension and by outcome (TALIS-PISA link average)



1. Unlike continuous variables where R^2 of linear regressions on school fixed effects are used to estimate the share of variance explained at the school level (right-part of the panel), in the case of binary variables analysed with the help of logistic regressions, an equivalent statistic to R^2 does not exist. The goodness-of-fit of logistic regression models is evaluated by pseudo R^2. Pseudo R^2 is on a similar scale as R^2 (ranging from 0 to 1) with higher values indicating better model fit. It is used as a proxy for the percentage of the variance in the dependent variable that the independent variables explain collectively.

Note: Teacher variables averaged for all teachers within the school.

Values are ranked in descending order of the percentage of variance in the index of classroom disciplinary climate explained at the school level.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 3.2, 3.4, 3.6 and 3.7.

StatLink https://doi.org/10.1787/888934223916

Do teacher and school factors matter equally across student social-emotional outcomes and countries?

This section aims to identify specific indicators of selected teacher and school dimensions that are significantly related with certain student outcomes, using country-level regressions. Only results that hold on average across the TALIS-PISA link countries and for, at least (and in some cases only), a couple of participating countries/economies, are reported. Regression analyses also provide insights into the potential role of confounding factors such as classmates' academic and socio-behavioural characteristics in the relationship between a teacher or school factor and student social-emotional development. Findings presented above may mask differences across the TALIS-PISA link countries and economies. In particular, relationships found on average across all nine participating countries and economies do not usually hold for every country/economy considered individually. This might be due to a lack of statistical power or simply reflect a true absence of relationship in the considered education system(s). This section therefore highlights the countries/economies in which the relationships identified are significant.

Teachers' classroom practices

The teacher dimension of classroom practices is consistently highlighted by the lasso regressions, variance decomposition analysis and education research as an important predictor of student social-emotional development. Country-level regression results further show that teachers' classroom practices are significantly related to student perception of the classroom climate and less so for the other three outcomes of interest in this chapter (Tables 3.9, 3.13, 3.17 and 3.19) (this is consistent with the results from both lasso regressions and variance decomposition analysis). First, there is a clear alignment of teachers' and students' perceptions of the classroom climate in all countries and economies. The more teachers report disciplinary issues, the more students report the same (though the relationship is not significant in Georgia and Turkey) (Table 3.9). Country-level regression results also show a mixed pattern regarding the use of cognitive activation (i.e. getting students to evaluate information and apply knowledge in order to solve a problem). The more teachers report using cognitive activation practices, the better the disciplinary climate reported by students in CABA (Argentina) and Turkey. The opposite is found for the Czech Republic and Denmark (for the latter, this holds true after accounting for the classroom composition only [Table 3.11]). While cognitive activation practices have been found to be associated with higher student performance in several countries (Echazarra et al., 2016[34]; Le Donné, Fraser and Bousquet, 2016[35]), this result might suggest that these practices are linked with student behaviour in class and overall with the classroom climate as they may generate more student agitation than other teaching strategies in some countries and less in others.

Classmates' characteristics and class size

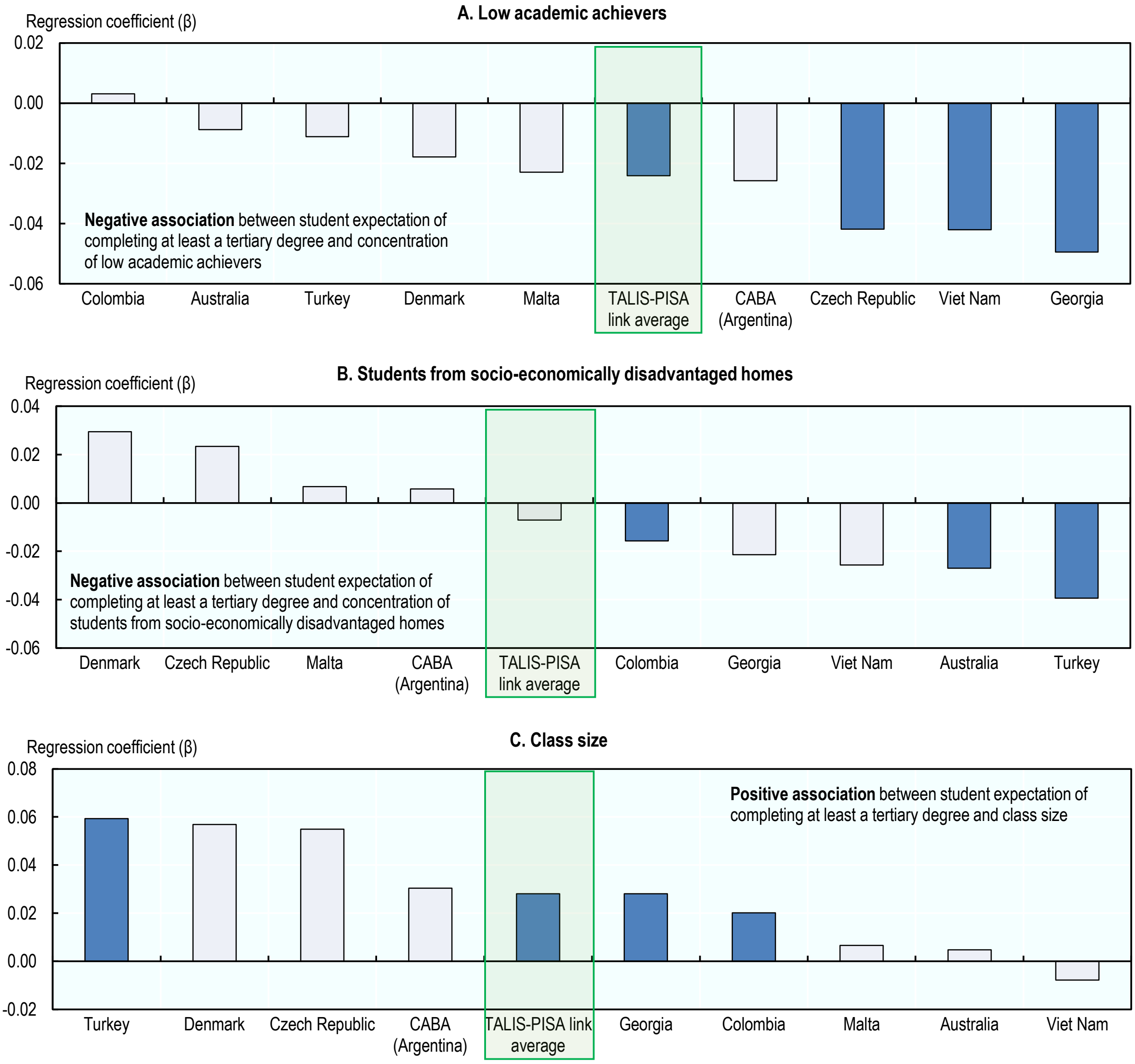
The classroom characteristics dimension explains a high share of school differences in student social-emotional outcomes, and the highest share for students' educational expectations (Figure 3.6). Lasso regression results identified classmates' characteristics, and in particular, the share of low academic achievers and students with behaviour problems as among the most important factors related to student social-emotional outcomes (Figure 3.5). Country-level regressions confirm these findings and reveal some other results.

There seems to be a relationship between the presence of high academic achievers in the class and student self-concept (Table 3.21). The more academically gifted students are enrolled in the classroom, the less difficult the student finds the PISA test. This holds true on average across countries and economies, for CABA (Argentina), the Czech Republic, Georgia and Turkey. This may be due to the fact that students attending high-achieving classrooms tend to be exposed to more complex testing situations.

In comparison to lasso results, country-level logistic regressions of student educational aspirations show that not only does the classroom presence of low academic achievers matter in the formation of students' educational expectations (Figure 3.7), but also that of students from socially disadvantaged homes. As the share of socio-economically disadvantaged students in the classroom increases, student expectation of attending tertiary education tends to decrease. This holds true in Australia, Colombia and Turkey. The fact that it remains true while controlling for students' own socio-economic background suggests important peer effects on student educational expectations in these countries. This might be because peers' aspirations influence a student's aspirations or because the student does not benefit from models of highly-educated adults among other peers' parents or, by extension, in the local community.

Figure 3.7. Relationship between classmates' characteristics and student educational expectations

Change in the expectation of completing at least a tertiary degree associated with classmates' characteristics



Notes: Results of logistic regressions based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status.
Teacher variables are averaged for all teachers within the school.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the expectation of completing at least a tertiary degree.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.22.

StatLink https://doi.org/10.1787/888934223935

Finally, these multiple regression analyses also highlight the potential role of class size in shaping both student self-concept and student educational aspirations (Tables 3.21 and 3.22). The larger the classes, the more favourable the outcomes (while accounting for classroom composition). The relationship with student perception of PISA test difficulty is significant in Denmark, Malta and Turkey and on average across countries. The relationship with educational aspirations is significant in Colombia, Georgia and Turkey and on average across countries. This suggests that the larger the peer group is, the better it is for student social-emotional outcomes. This might be due to the fact that the larger the class is, the more varied the family background, attitudes, behaviours and aspirations of other classmates are, and the richer and more favourable student social-emotional development is. Yet, most likely, the presence of a selection bias is also partly at play. A school's average class size partly captures unobserved characteristics about school intake, such as students' social-emotional skills. These fall outside those reported by teachers and which are accounted for in the regression (students' academic and socio-behavioural characteristics). It is common practice for schools and school systems to implement specific grouping strategies and adjust class sizes to the needs of their students. A previous PISA report found there was a significant difference of more than three students per class between socio-economically advantaged and disadvantaged schools on average across the OECD, and that such a positive and significant difference was found in 39 education systems (OECD, 2018, p. 89[36]).

School culture and school leadership

Multiple linear regressions show that positive teacher-student relations typically matter for two classroom-related outcomes: student perception of classroom disciplinary climate and teacher enthusiasm (Tables 3.23 and 3.27). The more teachers report getting along with their students, the more disciplined the classrooms are on average across countries and economies, and in Australia, Colombia, Georgia and Turkey (and in Malta when focusing on reading teachers only [Table 3.24]) according to students. After controlling for classroom characteristics, the relationship remains significant in Colombia and Georgia (Table 3.25). Likewise, the more teachers report nurturing good relationships with students, the more students perceive them as enjoying teaching. This holds true on average across participating countries and economies, in Australia, Colombia, Denmark, Georgia and Turkey. It remains significant in Australia, Colombia and Georgia after accounting for classroom characteristics (Table 3.29).

With regard to school leadership, the more frequently principals report collaborating with teachers to solve classroom discipline problems, the less disciplined the classroom climate is as reported by students. This relationship is found to be significant in Colombia, the Czech Republic, Denmark, Georgia and Malta either when focusing on all teachers of the school or on reading teachers only (Tables 3.31 and 3.32). It tends to remain significant after accounting for classroom composition (Tables 3.33 and 3.34).

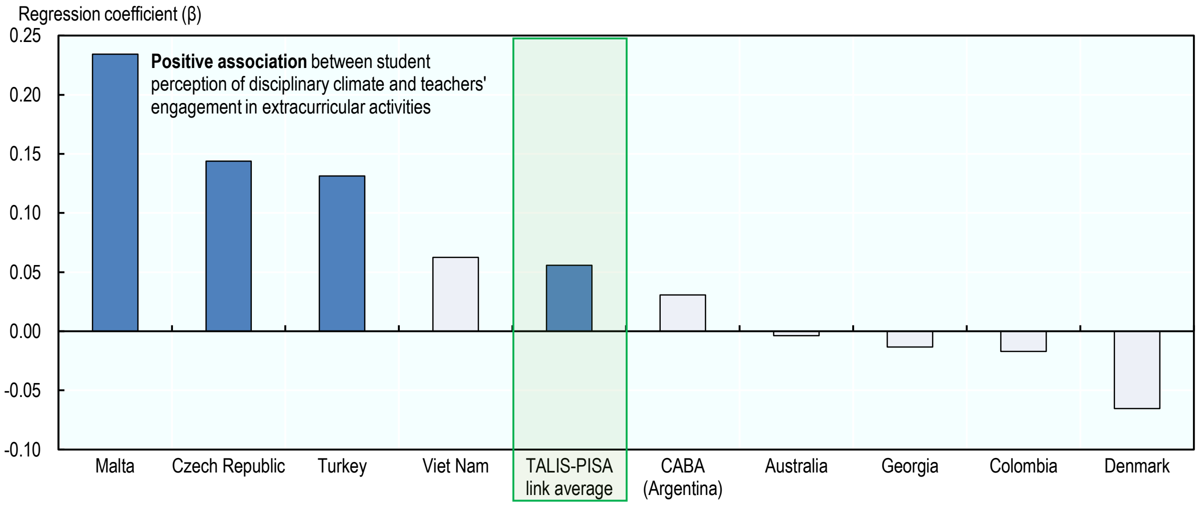
Teachers' use of working time

The way teachers use and allocate their working time across various professional activities is related to the way students perceive their learning environment and shape their educational expectations. In particular, the amount of time teachers spend on extracurricular activities is positively related to student perceptions of classroom disciplinary climate and teacher enthusiasm. It is also positively related to students' educational expectations, at least in a couple of countries (Tables 3.35, 3.39 and 3.43). This is particularly important for classroom disciplinary climate. The more teachers spend time with students on extracurricular activities, the better the classroom disciplinary climate on average across countries, in the Czech Republic, Malta, Turkey and Viet Nam. After accounting for classroom composition, the relationship remains significant except in Viet Nam (Figure 3.8). Extracurricular activities at school usually aim to achieve a broad set of goals, such as physical exercise and health, the development of creativity and practice or appreciation of the arts, or volunteering and engagement with the community (OECD, 2020, p. 146[30]). Past research found that participation in extracurricular activities can help students develop social-emotional skills, such as persistence, teamwork or a stronger sense of belonging at school. A literature review of studies conducted on extracurricular activities in the early 2000s looked at the relationships between adolescents' participation in extracurricular activities and their academic achievement, substance use, sexual activity, psychological adjustment, and delinquency (Farb and

Matjasko, 2012[37]). Yet, this meta-analysis did not emphasise the importance of teachers as important moderators in these relationships. The finding from this report suggests that having the teachers themselves spend time with students outside of the usual lessons is beneficial to the classroom climate and possibly to the overall learning environment and development of students.

Figure 3.8. Relationship between teachers' engagement in extracurricular activities and student perception of classroom disciplinary climate

Change in the index of student perception of the classroom disciplinary climate associated with teachers' engagement in extracurricular activities



Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status.
Teacher variables are averaged for all teachers within the school.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in the index of student perception of the classroom disciplinary climate.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.37.

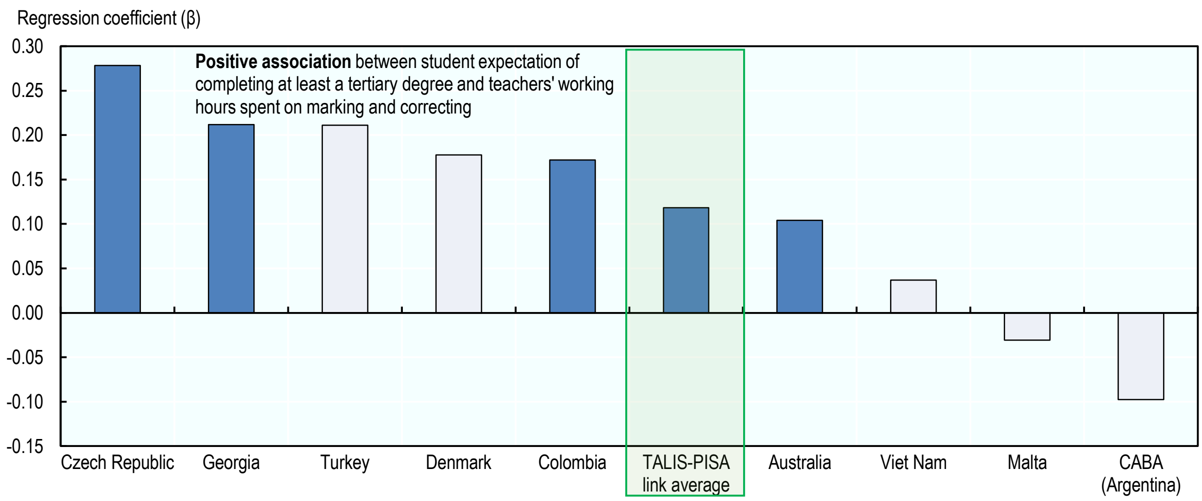
StatLink https://doi.org/10.1787/888934223954

Another teacher activity that matters for students' self-concept and formation of educational expectations is the marking and correcting of student work (Figure 3.9). The more hours teachers report spending on this per week, the less difficult students find the PISA test and the more likely they expect to complete at least a tertiary degree. Together with the positive relationship found in Chapter 2 with student achievement, this finding points to the key role teachers' engagement plays in evaluating, reflecting and communicating on student progress in student academic success and educational expectations. This is consistent with findings from past research highlighting the "power of feedback" (Hattie and Timperley, 2007[38]; Wisniewski, Zierer and Hattie, 2020[39]). Results from a recent meta-analysis (Wisniewski, Zierer and Hattie, 2020[39]) indicate a medium effect of feedback on student learning. Yet, the impact of feedback is substantially influenced by what is conveyed. Furthermore, feedback has higher impact on cognitive skills outcomes than motivational and behavioural outcomes.

Like extracurricular activities, marking and correcting student work in a constructive and efficient manner seems to be a valuable way for teachers to support student development. Yet, as commented in Chapter 2, the relationship may run in the opposite direction: that it is because students have higher expectations and greater self-concept in some schools than in others that teachers in these schools invest more time in assessing and marking these students in order to meet their expectations.

Figure 3.9. Relationship between time spent by teachers on marking and correcting student work and student educational expectations

Change in the expectation of completing at least a tertiary degree associated with time spent by teachers on marking and correcting student work



Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status.
Teacher variables are averaged for all teachers within the school.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in the expectation of completing at least a tertiary degree.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.44.

StatLink https://doi.org/10.1787/888934223973

Other teacher activities are related to student perception of their learning environment and educational aspirations, yet not in a favourable way, suggesting a reverse causality. For example, the less disciplined a classroom is, the more time teachers report spending on communication and co-operation with parents on average across countries but also in the Czech Republic, Turkey and Viet Nam (and Colombia after controlling for classroom composition) (Table 3.35). Also, the lower student educational expectations are, the more teachers report spending time on team work and dialogue with colleagues in their school on average across countries and in Georgia and Turkey (and the Czech Republic after accounting for classroom composition) (Table 3.43). Similarly, the less likely students are to aspire to complete a tertiary education degree, the more often teachers participate in school management on average across countries and in Australia, Turkey and Viet Nam (and CABA [Argentina] after controlling for classroom composition). For a few countries, these relationships disappear after controlling for classroom composition, which suggests that the student's classroom characteristics mediate the relationship between the student's aspiration and teachers' engagement in school management and dialogue with families (Tables 3.37 and 3.44). This does not necessarily mean that communicating with parents and colleagues is detrimental to student social-emotional development. Rather, it likely means that teachers spend more time doing this when they feel students need it.

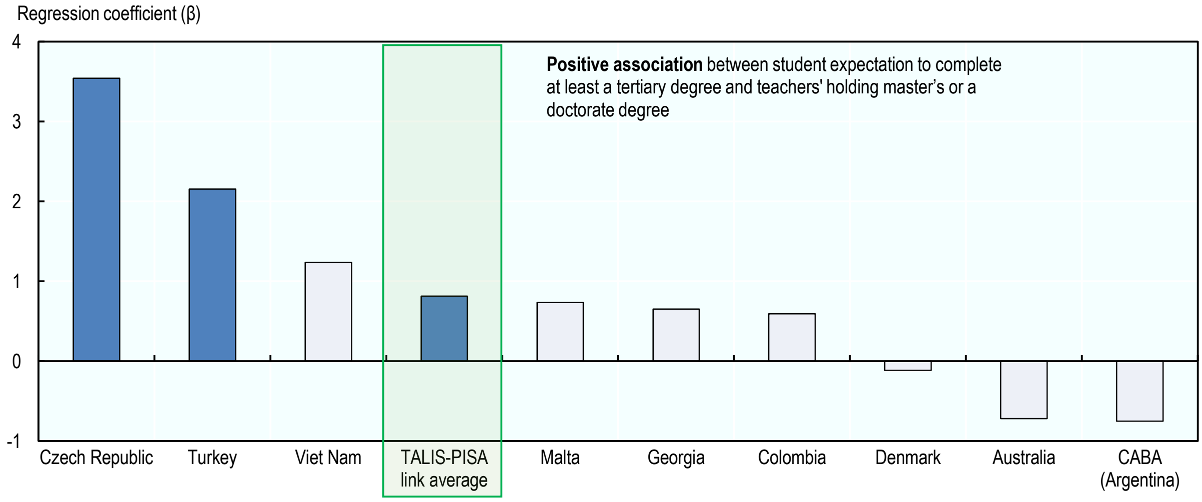
Teacher initial education

The level and content of teachers' formal education seem to matter for students' perceptions of their learning environment and their own educational expectations. One original finding is the positive relationship between the share of teachers holding a master's or a doctorate degree in the school and student expectation of completing at least a tertiary degree (Table 3.45). The greater the number of teachers who hold a master's or a doctorate degree at the school, the more likely students are to expect completing a tertiary degree on average across countries, as well as in the Czech Republic and Turkey. This remains true even after accounting for classroom socio-demographic and academic composition (as reported by the teachers) (Figure 3.10), suggesting that there is a net "effect" highly-educated teachers have on the way students shape their educational aspirations. This may be because these teachers are better informed about the existing tracks toward completing a tertiary degree, or because they embody students' higher educational aspirations. In addition, the more teachers felt prepared for at least some, if not all, the subject(s) they teach, the higher the educational expectations of the students. This holds true on average across countries, as well as in the Czech Republic, Georgia, Malta and Viet Nam, but only remains significant for Georgia and Viet Nam after controlling for classroom socio-demographic and academic composition (Tables 3.47 and 3.48).

These results show that teachers' levels of education (and perceived levels of preparation) may be related to student educational aspirations in a couple of the nine participating education systems. This adds to the debate and the lack of consensus on the effect of teacher education on student achievement (see Chapter 2 on this). While Rivkin, Hanushek and Kain (2005[40]) report no clear evidence that a master's degree improves teacher effectiveness on student academic achievement, the TALIS-PISA link data analysis suggests that a master's or a doctoral degree and appropriate teacher initial preparation may make a difference for different student outcomes, namely student educational aspirations. Yet, once again and despite the controls introduced in the analyses to account for classroom composition and student personal characteristics, the direction of the relationship could potentially be the reverse. Teachers' credentials and certification can influence teachers' employment conditions and, in particular, school assignment (OECD, 2018[36]).[13] Teachers with more education might work in different schools, either because they are given more choice in school assignment or because education authorities allocate teachers to different schools based on their qualifications. Teacher assignment to schools might well be driven by teacher education level as well as the academic level and socio-demographic composition of the school and the social-emotional skills of its students or a combination of the above. In education systems offering several educational programmes to students from an early age, such as the Czech Republic and Turkey (OECD, (2020[30]), Figure V.3.2), more educated teachers are more likely to teach in the more academic and general tracks with students who have higher educational expectations. All this could explain the positive relationship found between teacher educational level and student educational expectations observed in the Czech Republic and Turkey.

Figure 3.10. Relationship between teacher educational level and student educational expectations

Change in the expectation of completing at least a tertiary degree associated with teacher educational level



Notes: Results of a logistic regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status.
Teacher variables are averaged for all teachers within the school.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in the expectation of completing at least a tertiary degree.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.46.

StatLink https://doi.org/10.1787/888934223992

Teachers' participation in professional development

There is no one single feature of teacher professional development activities that seems to matter for more than one of the four student social-emotional outcomes under scrutiny. Each outcome is related to specific features of professional development activities. Having teachers physically attend courses/seminars in a school seems beneficial to classroom disciplinary climate as perceived by the students (Table 3.50). This holds on average across participating countries but also in Australia, Denmark and Turkey and remains true, as well as becomes true, for Malta after accounting for classroom composition (Table 3.52). The content of teacher professional development activities matters for students' perception of a teacher's enthusiasm. It is associated as well with students' educational expectations of completing a tertiary degree (Tables 3.53 and 3.57). The more teachers in a school participate in professional development focusing on the knowledge and understanding of their subject fields, the more students feel that their teachers like and enjoy teaching. This positive relationship is significant on average across participating countries, in the Czech Republic, Malta and Viet Nam, and remains significant even after controlling for classroom composition (Table 3.55).

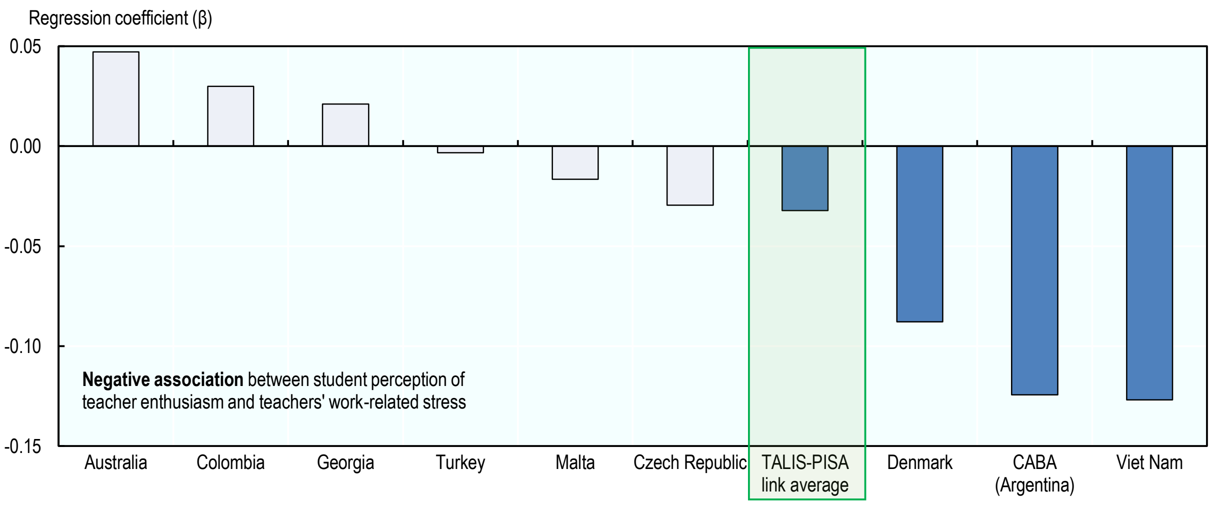
Teachers' well-being and job satisfaction

Each of the four student social-emotional outcomes is related in some way to teachers' own social-emotional well-being at work. Students tend to find their teachers more interested in teaching when those teachers report lower levels of work-related stress on average in their school. This holds true in

CABA (Argentina), Denmark and Viet Nam, and on average across all countries and economies when focusing on reading teachers only (this is also consistent with results from the lasso regression on the pooled sample) (Table 3.64). This relation remains generally significant after accounting for classroom composition (Figure 3.11).

Figure 3.11. Relationship between teacher work-related stress and student perception of teacher enthusiasm

Change in the index of student perception of teacher enthusiasm associated with work-related stress



Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status.
Teacher variables are averaged only for subject domain teachers. The analysis is based on a sample restricted to schools with at least one subject domain teacher.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in the index of student perception of teacher enthusiasm.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 3.66.

StatLink https://doi.org/10.1787/888934224011

Teachers' satisfaction with their work environment seems to be a positive factor for two student social-emotional outcomes: student behaviour in class as measured by classroom disciplinary climate (the relationship is significant on average across countries and economies, in Australia, CABA [Argentina] and Turkey) and with students' feeling of mastery over the PISA test (the relation is significant on average across countries, in Australia, Colombia, Georgia and Turkey) (Tables 3.59 and 3.67). Yet, these relationships vanish in almost all countries once classroom characteristics have been accounted for (Tables 3.61 and 3.68).

The TALIS-PISA link data also indicate some relation between teachers' satisfaction with salary and student educational expectations. On average across countries and economies as well as in Malta, Turkey and Viet Nam (and in the Czech Republic after accounting for classroom composition), the less satisfied teachers are with their salary in the school, the more likely students are to contemplate completing at least a tertiary degree (Tables 3.69 and 3.70). Under the same conditions on average across countries and economies as well as in Malta and Turkey, students also feel more able to master the PISA test (Tables

3.67 and 3.68). This suggests that some teachers not only project higher expectations for themselves in terms of salary level but also influence the educational expectations and confidence of their students upward. This influence may be direct when teachers are very explicit about their expectations vis-à-vis students or suggestive when teachers implicitly transmit their beliefs about valuable jobs and studies to students. The fact that this relationship remains significant after controlling for classroom composition in all concerned countries but Malta suggests that it is not fully mediated by the influence of students' classmates. Colombia is the exception to this pattern: the more satisfied teachers are with their salary, the higher students' educational expectations. Indicators from *Education at a Glance* suggest that this relationship is mediated by teacher experience as Colombian teachers receive a substantial increase in salary after 10 years of experience (one of the sharpest worldwide) and teachers who reach the top of the salary scale can earn more than double the starting salary (OECD, 2020[41]). However, further analysis of Colombia's TALIS-PISA link data shows that the relationship remains significantly positive after controlling for both teacher experience and classroom composition.

Summary

Three main conclusions can be drawn from the findings presented above.

First, there seem to be clear school patterns of disciplinary climate. There is a high degree of alignment in different stakeholders' perceptions of school climate. And there is consistency in the measures taken by school leaders and teachers to address disciplinary issues in the school. Teachers' reporting of disciplinary issues is matched by students reporting of the same in all countries/economies with the exception of Georgia and Turkey where the relationship is not significant. Likewise, the less disciplined the classroom climate is, as reported by students, the more frequently principals report collaborating with teachers to solve classroom discipline problems, and the more time teachers report spending on communication and co-operation with parents. This does not necessarily mean that principals' and teachers' remedial measures are detrimental to student behaviour in class. Rather, principals and teachers invest more time in these kinds of activities when they feel students are in need of it. Some findings suggest that such vicious circles could be inverted by establishing positive relationships between teachers and students, and promoting a favourable climate for teachers' work well-being. Indeed, results show that the more teachers report nurturing good relationships with students, the more students perceive them as enjoying teaching and the better it is for the disciplinary climate. In addition, students tend to find their teachers more interested in their teaching when teachers report lower levels of work-related stress on average in the school.

Second, the characteristics and number of classmates matter for student social-emotional skills, in particular for student self-concept. The greater the number of academically gifted students enrolled in the classroom, the more students feel able to succeed in the PISA test. As well, as the share of socio-economically disadvantaged students in the classroom increases, student expectation of completing tertiary studies tends to decrease. These findings suggest the positive effect of increased exposure to high-performing peers on students' immediate self-concept, and a greater sensitivity of longer-term self-concept to the presence of socially disadvantaged peers. On average, students might be more easily influenced by the lower aspirations of their socially disadvantaged peers or by the lack of highly educated adults as role models among their peers' parents or, by extension, in the local community.

Third, although causality cannot be established, two teacher characteristics and practices have been highlighted as positively related to student social-emotional development: the level and quality of teacher pre-service training, and the way teachers use their working time. Results show that the greater the number of teachers who hold a master's or a doctorate degree there are in the school, and the more they felt prepared to teach their subject after their initial training, the more likely students are to expect to complete at least a tertiary degree. Then, as the amount of time teachers spend on extracurricular activities increases, the more students report that the classroom is disciplined, that the teacher is interested and motivated to teach and that they expect to complete at least a tertiary degree. In addition, the more time teachers spend on marking and correcting student work, the higher the level of student self-concept of doing well in the PISA test and completing tertiary studies. These findings suggest that spending quality time with students outside of the usual lessons and devoting time to assess student work and communicate progress to students are valuable ways to support student social-emotional development.

References

Avvisati, F., N. Le Donné and M. Paccagnella (2019), "A meeting report: cross-cultural comparability of questionnaire measures in large-scale international surveys", *Measurement Instruments for the Social Sciences*, Vol. 1/8, pp. 1-10, http://dx.doi.org/10.1186/s42409-019-0010-z. [45]

Barratt, E. (1995), "History of personality and intelligence theory and research", in Saklofske, D. and M. Zeidner (eds.), *International Handbook of Personality and Intelligence. Perspectives on Individual Differences*, Springer, Boston, MA, http://dx.doi.org/10.1007/978-1-4757-5571-8_1. [6]

Blazar, D. et al. (2017), "Attending to general and mathematics-specific dimensions of teaching: Exploring factors across two observation instruments", *Educational Assessment*, Vol. 22/2, pp. 71-94, http://dx.doi.org/10.1080/10627197.2017.1309274. [22]

Blazar, D. and M. Kraft (2017), "Teacher and teaching effects on students' attitudes and behaviors", *Educational Evaluation and Policy Analysis*, Vol. 39/1, pp. 146-170, http://dx.doi.org/10.3102/0162373716670260. [12]

Borghans, L. et al. (2008), "The economics and psychology of personality traits", *Journal of Human Resources*, Vol. 43/4, pp. 972-1059, http://dx.doi.org/10.1353/jhr.2008.0017. [9]

Borgonovii, F. and J. Pál (2016), "A Framework for the Analysis of Student Well-Being in the PISA 2015 Study: Being 15 In 2015", *OECD Education Working Papers*, No. 140, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jlpszwghvvb-en. [44]

Breda, T. et al. (2020), "Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools", *PSE Working Papers*, No. 2018-06, HAL Sciences de l'Homme et de la Société, https://halshs.archives-ouvertes.fr/halshs-01713068v3. [27]

Carrell, S. and M. Hoekstra (2010), "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids", *American Economic Journal: Applied Economics*, Vol. 2/1, pp. 211-228, http://dx.doi.org/10.1257/app.2.1.211. [32]

Carrell, S., M. Page and J. West (2010), "Sex and science: How professor gender perpetuates the gender gap", *The Quarterly Journal of Economics*, Vol. 125/3, pp. 1101-1144, http://dx.doi.org/10.1162/qjec.2010.125.3.1101. [28]

Chetty, R. et al. (2011), "How does your kindergarten classroom affect your earnings? Evidence from Project STAR", *The Quarterly Journal of Economics*, Vol. 126/4, pp. 1593-1660, http://dx.doi.org/10.1093/qje/qjr041. [10]

Chetty, R., J. Friedman and J. Rockoff (2014), “Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood”, *American Economic Review*, Vol. 104/9, pp. 2633-2679, http://dx.doi.org/10.1257/aer.104.9.2633. [14]

Duckworth, A. et al. (2007), “Grit: Perseverance and passion for long-term goals”, *Journal of Personality and Social Psychology*, Vol. 92/6, pp. 1087-1101, http://dx.doi.org/10.1037/0022-3514.92.6.1087. [7]

Duckworth, A., P. Quinn and E. Tsukayama (2012), “What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades”, *Journal of Educational Psychology*, Vol. 104/2, pp. 439-451, http://dx.doi.org/10.1037/a0026280. [8]

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. [34]

Farb, A. and J. Matjasko (2012), “Recent advances in research on school-based extracurricular activities and adolescent development”, *Developmental Review*, Vol. 32/1, pp. 1-48, http://dx.doi.org/10.1016/j.dr.2011.10.001. [37]

Gaviria, A. and S. Raphael (2001), “School-based peer effects and juvenile behavior”, *Review of Economics and Statistics*, Vol. 83/2, pp. 257-268, http://dx.doi.org/10.1162/00346530151143798. [33]

Gershenson, S. (2016), “Linking teacher quality, student attendance, and student achievement”, *Education Finance and Policy*, Vol. 11/2, pp. 125-149, http://dx.doi.org/10.1162/EDFP_a_00180. [15]

Hafen, C. et al. (2015), “Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system-secondary”, *The Journal of Early Adolescence*, Vol. 35/5-6, pp. 651-680, http://dx.doi.org/10.1177/0272431614537117. [23]

Hanushek, E. and S. Rivkin (2010), “Generalizations about using value-added measures of teacher quality”, *American Economic Review*, Vol. 100/2, pp. 267-271, http://dx.doi.org/10.1257/aer.100.2.267. [4]

Hattie, J. and H. Timperley (2007), “The power of feedback”, *Review of Educational Research*, Vol. 77/1, pp. 81-112, http://dx.doi.org/10.3102/003465430298487. [38]

Hoy, W., J. Hannum and M. Tschannen-Moran (1998), “Organizational climate and student achievement: A parsimonious and longitudinal view”, *Journal of School Leadership*, Vol. 8/4, pp. 336-359, http://dx.doi.org/10.1177/105268469800800401. [3]

Jackson, C. (2018), “What do test scores miss? The importance of teacher effects on non–test score outcomes”, *Journal of Political Economy*, Vol. 126/5, pp. 2072-2107, http://dx.doi.org/10.1086/699018. [16]

Jackson, C. (2012), “Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina”*, NBER Working Paper Series*, No. 18624, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w18624. [26]

Jennings, J. and T. DiPrete (2010), “Teacher effects on social and behavioral skills in early elementary school”, *Sociology of Education*, Vol. 83/2, pp. 135-159, http://dx.doi.org/10.1177/0038040710368011. [17]

Kankaraš, M. and J. Suarez-Alvarez (2019), “Assessment framework of the OECD Study on Social and Emotional Skills”, *OECD Education Working Papers*, No. 207, OECD Publishing, Paris, https://dx.doi.org/10.1787/5007adef-en. [43]

Koedel, C. (2008), “Teacher quality and dropout outcomes in a large, urban school district”, *Journal of Urban Economics*, Vol. 64/3, pp. 560-572, http://dx.doi.org/10.1016/j.jue.2008.06.004. [18]

Kraft, M. (2019), “Teacher effects on complex cognitive skills and social-emotional competencies”, *Journal of Human Resources*, Vol. 54/1, pp. 1-36, http://dx.doi.org/10.3368/jhr.54.1.0916.8265R3. [19]

Ladd, H. and L. Sorensen (2017), “Returns to teacher experience: Student achievement and motivation in middle school”, *Education Finance and Policy*, Vol. 12/2, pp. 241-279, http://dx.doi.org/10.1162/EDFP_a_00194. [20]

Le Donné, N., P. Fraser and G. Bousquet (2016), “Teaching Strategies for Instructional Quality: Insights from the TALIS-PISA Link Data”, *OECD Education Working Papers*, No. 148, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jln1hlsr0lr-en. [35]

Moffitt, T. et al. (2011), “A gradient of childhood self-control predicts health, wealth, and public safety”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 108/7, pp. 2693-2698, http://dx.doi.org/10.1073/pnas.1010076108. [11]

OECD (2020), *Education at a Glance 2020: OECD Indicators*, OECD Publishing, Paris, https://dx.doi.org/10.1787/69096873-en. [41]

OECD (2020), *Global Teaching InSights: A Video Study of Teaching*, OECD Publishing, Paris, https://dx.doi.org/10.1787/20d6f36b-en. [24]

OECD (2020), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/ca768d40-en. [30]

OECD (2020), *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/19cf08df-en. [42]

OECD (2019), *PISA 2018 Results (Volume III): What School Life Means for Students’ Lives*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/acd78851-en. [1]

OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/1d0bc92a-en. [2]

OECD (2018), *Effective Teacher Policies: Insights from PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264301603-en. [36]

OECD (2015), *Skills for Social Progress: The Power of Social and Emotional Skills*, OECD Skills Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264226159-en. [25]

Pianta, R. and B. Hamre (2009), “Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity”, *Educational Researcher*, Vol. 38/2, pp. 109-119, http://dx.doi.org/10.3102/0013189X09332374. [13]

Rivkin, S., E. Hanushek and J. Kain (2005), “Teachers, schools, and academic achievement”, *Econometrica: Journal of the Econometric Society*, Vol. 73/2, pp. 417-458, http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x. [40]

Ruzek, E. et al. (2015), “Using value-added models to measure teacher effects on students’ motivation and achievement”, *The Journal of Early Adolescence*, Vol. 35/5-6, pp. 852-882, http://dx.doi.org/10.1177/0272431614525260. [21]

Sacerdote, B. (2011), “Peer effects in education: How might they work, how big are they and how much do we know thus far?”, in Hanushek, E., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1. [31]

Todd, P. and K. Wolpin (2007), “The production of cognitive achievement in children: Home, school, and racial test score gaps”, *Journal of Human Capital*, Vol. 1/1, pp. 91-136, http://dx.doi.org/10.1086/526401. [5]

Van de gaer, E. et al. (2012), “The reference group effect: An explanation of the paradoxical relationship between academic achievement and self-confidence across countries”, *Journal of Cross-Cultural Psychology*, Vol. 43/8, pp. 1205-1228, http://dx.doi.org/10.1177/0022022111428083. [29]

Wisniewski, B., K. Zierer and J. Hattie (2020), “The power of feedback revisited: A meta-analysis of educational feedback research”, *Frontiers in Psychology*, Vol. 10/3087, pp. 1-14, http://dx.doi.org/10.3389/FPSYG.2019.03087. [39]

Notes

[1] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[2] For instance, analysis based on TALIS data show that, in the case of teachers’ satisfaction with the profession, only a small percentage (i.e. 4%) of the total variance comes from differences between schools (OECD, 2020[42]).

[3] The terms “social and emotional outcomes” and “social and emotional development” are used in continuity with the terminology employed in other OECD work – see, in particular, Box I.1 of the OECD Working Paper by Kankaraš and Suarez-Alvarez (2019[43]).

[4] PISA measures classroom disciplinary climate by asking students how frequently ("never or hardly ever", "some lessons", "most lessons", "every lesson") the following things happen in their language-of-instruction lessons: "Students don't listen to what the teacher says"; "There is noise and disorder"; "The teacher has to wait a long time for students to quiet down"; "Students cannot work well"; and "Students don't start working for a long time after the lesson begins". These statements were combined to create the index of classroom disciplinary climate whose average is 0 and standard deviation is 1 across OECD countries. Positive values on this scale mean that the student enjoys a better disciplinary climate in language-of-instruction lessons than the average student in OECD countries.

[5] PISA measures student perception of teacher enthusiasm in language-of-instruction lessons. PISA asked 15-year-old students whether they agree ("strongly disagree", "disagree", "agree", "strongly agree") with the following statements about the teacher teaching the two language-of-instruction lessons they attended prior to sitting the PISA test: "It was clear to me that the teacher liked teaching us"; "The enthusiasm of the teacher inspired me"; "It was clear that the teacher likes to deal with the topic of the lesson"; and "The teacher showed enjoyment in teaching". These statements were combined to create the index of teacher enthusiasm whose average is 0 and standard deviation is 1 across OECD countries. Positive values in this index mean that students perceived their language-of-instruction teachers to be more enthusiastic than the average student across OECD countries did.

[6] The original PISA 2015 framework for student well-being was more comprehensive than the one retained in PISA 2018 and in this report. Beyond the cognitive and psychological dimensions of well-being, it included three additional dimensions: social, physical and material well-being (Borgonovii and Pál, 2016[44]).

[7] PISA measures perception of difficulty of the PISA test by asking students how they feel about the reading tasks included in the test ("strongly agree", "disagree", "agree", "strongly agree"): "There were many words I could not understand"; "Many texts were too difficult for me"; "I was lost when I had to navigate between different pages". These statements were combined to create the index of perception of difficulty of the PISA test whose average is 0 and standard deviation is 1 across OECD countries. Positive values on this scale mean that the student finds the test more difficult.

[8] Alternative coding was tried but the one discriminating expectations of completing at least a tertiary degree versus lower educational degrees was found to be subject to the largest differences between schools. Therefore, it was retained for the following analyses.

[9] Unlike the other student outcomes, student educational expectation is coded as a binary variable. Therefore, the share of variance between schools cannot be estimated as it is for the continuous variables. The pseudo-R^2 of the logistic regression model, including dummies for all participating schools, is used as a proxy (see Annex B for more information).

[10] In Georgia, schools show differences of similar magnitude in the other indicators of student interest in school.

[11] Longer questions, including many sub-items, seem also to be a better option to enhance cross-cultural comparability of measurement (Avvisati, Le Donné and Paccagnella, 2019[45]).

[12] This result is based on lasso regressions in which many predictors of student social-emotional development are included. This means that, among many other covariates, the share of low academic achievers and the share of students with behavioural problems are jointly introduced into the regressions.

[13] An OECD report based on PISA 2018 data showed that teachers are sorted across schools with different socio-economic backgrounds according to their qualification level. More specifically, on average across OECD countries, only 69% of science teachers in disadvantaged schools have a university degree with a major in science while 79% of science teachers in advantaged schools did. This makes for a significant difference of 10 percentage points between the top and bottom quarters of school socio-economic profiles (OECD, 2018, p. 97[36]).

4 Does what teachers and schools do matter differently depending on student achievement and gender?

Providing equal educational opportunities means that all students, irrespective of their gender or background, have the same chances of fulfilling their potential. However, despite significant efforts made by societies to narrow disparities in students' outcomes in the recent past, gaps still persist. Drawing on the rich TALIS-PISA link data, this chapter explores whether certain teacher and school factors that are identified in Chapter 2 of this report as key predictors of student achievement for average performing students also matter for low achievers and their high-achieving peers. In addition, the chapter investigates the teacher and school factors that are significantly related to within-school disparities in performance between girls and boys.

Highlights

- On average across the countries and economies participating in the TALIS-PISA[1] link, certain teacher and school factors, including teachers' class time spent on actual teaching and learning, teachers' satisfaction with the work environment, teachers' working hours spent on marking and correcting, the concentration of academically gifted students in the classroom as well as the degree of parents' and community's involvement in school-related activities, matter both for low and high-achieving students.
- On average across the countries and economies participating in the TALIS-PISA link, the share of teachers reporting workload as an important source of stress at the school may matter specifically for low-performing students in relation to their performance in reading and mathematics. Teachers' workload stress is significantly associated with better performances for low-achieving students in both subject domains. This may signal teachers' commitment and dedication to their work in general. More concretely, it may hint at teachers' extra effort in helping low-achieving students, which, in turn, can lead to additional workload-induced stress.
- Girls outperform boys in reading in around one-third of the schools, on average across the countries and economies participating in the TALIS-PISA link. The opposite pattern is observed for mathematics and science, with boys performing better than girls, although in fewer schools, unlike girls' overperformance in reading. In quite a few schools in each participating country and economy, boys perform equally well or better than girls in reading and girls perform equally well or better than boys in mathematics and science. This suggests that gender gaps can be closed within schools.
- As the level of disciplinary issues that the average school teacher perceives in the classroom increases, the outperformance of girls over boys in reading in the school tends to increase on average across the countries and economies participating in the TALIS-PISA link. This signals that boys might be more affected by a deteriorated classroom climate as they tend to be exposed to greater peer pressure than girls, but also as they may be more likely to disturb classes.
- On average across the countries and economies participating in the TALIS-PISA link, girls tend to outperform boys in reading even more in those schools where principals report more often that they observe instruction in the classroom and resolve problems with the lesson timetable. On the contrary, boys' performance increases compared to that of girls when the school leader more regularly ensures that teachers feel responsible for their students' learning outcomes.
- Teacher-student relations as well as stakeholder involvement in school seem to matter for school-level gender differences in reading performance. In schools where the average school teacher considers teacher-student relationships to be positive, disparities in reading performance between girls and boys tend to decrease in favour of boys on average across countries and economies participating in the TALIS-PISA link. Similarly, in Colombia and Denmark, boys are more likely to perform as well as (if not better than) girls in reading in schools where the degree of parents' and community's involvement in school-related activities is higher.
- Teachers may have greater influence on the academic performance of students of their gender either because they apply distinct teaching practices that suit same-gender students better or because they address gender stereotypes by acting as role models for students of their gender. In Ciudad Autónoma de Buenos Aires (hereafter CABA [Argentina]), the Czech Republic and Turkey, the differences in performance in reading, mathematics and science between girls and boys tend to change in favour of girls as the share of female teachers increases in the school. In the same way, as the share of male teachers increases in the school, the better boys perform compared to girls in all subjects.

Introduction

Succeeding in today's fast-changing world requires skills rooted in academic competences such as literacy, numeracy and science but also the ability to think critically, cope with unforeseen problems, communicate effectively and work in teams. Teaching market-relevant skills to all students is more important than ever. Poor or inadequate skills limit access to better-paying and more rewarding jobs and, more generally, to better living and health conditions and higher social and political participation (Hanushek et al., 2015[1]; OECD, 2016[2]). Extensive research confirms the impact of low academic performance on future educational and socio-economic development – see, for example, Erikson et al. (2005[3]) and Rose and Betts (2004[4]).

Schools and education systems have the ambition to provide equal learning opportunities to all students. Equity does not mean that all students must obtain equal outcomes but rather that, provided with the same opportunities, differences in students' outcomes are not driven by individual factors like gender, socio-economic status, immigration background or disabilities (OECD, 2018[5]) – on the equality of opportunity theory see, for example, the influential contribution by Roemer (1998[6]). In other words, giving equal opportunities to all students means that, boys and girls, socially advantaged and socially disadvantaged students as well as native and migrant students have the same chances of fulfilling their potential. Giving equal opportunities to all students also means that the implemented teaching practices and teacher-related policies benefit as many students as possible without being detrimental to a few – either the most fragile or the most talented – and that they strike the collective optimum.

However, despite significant efforts made by societies to narrow disparities in students' outcomes in the recent past, gaps still persist (OECD, 2019[7]; OECD, 2018[5]). According to PISA results, more boys than girls score below the baseline level of proficiency in reading, mathematics and science; among high-performing students there are fewer girls than boys in areas such as mathematics, science and problem solving. Yet, even when boys and girls are equally proficient, their attitudes towards learning and aspirations are very different (OECD, 2019[7]; OECD, 2015[8]). Students' family socio-economic status is widely recognised as a reliable predictor of their academic performance and, indirectly, success in life. Analyses based on PISA data show consistently that, while many disadvantaged students succeed at school, advantaged students tend to outperform their disadvantaged peers in all subjects. Disparities in outcomes by students' immigration background have been observed in all PISA cycles. Generally, non-immigrant students outperform their first- and second-generation immigrant peers even after accounting for students' and schools' socio-economic profile (OECD, 2019[7]). Nevertheless, this association is closely related to the socio-economic background of immigrant students, and thus may not hold for countries that apply selective immigration policies favouring highly-skilled immigrants.

This chapter explores the characteristics and practices of schools and their teachers that matter for equity in student achievement. It draws on the rich TALIS-PISA link 2018 dataset of 15-year-old students, their teachers and schools from eight countries and economies[2] – Australia, CABA (Argentina), Colombia, the Czech Republic, Denmark, Georgia, Malta and Turkey. It addresses the following research questions: Do teacher and school factors matter equally for students with different academic performance? How can schools mitigate gender disparities in student performance? To address the first aforementioned research question, the chapter applies quantile regressions to explore whether the factors identified in Chapter 2 (*What do teachers and schools do that matters most for student achievement?)* as key predictors of PISA test scores for average performing students also matter for low achievers and their high-achieving peers. The first section also discusses the practices and characteristics of teachers and schools that matter specifically for students with different academic performances. The next section investigates the teacher and school factors that are significantly related to within-school disparities in performance between girls and boys. It focuses on gaps among the students within schools rather than on gaps at the country level, which are much widely analysed and commented in general. While it would have been interesting and relevant from a policy viewpoint to analyse gaps in achievement between socially advantaged students and socially disadvantaged students[3] as well as between native and migrant students[4] within schools, it was decided not to pursue this path for technical and analytical reasons. Nevertheless, since academic achievement and socio-economic status are positively related in all countries (OECD, 2019[7]), exploring the teacher and school factors that matter specifically for low achievers can be also informative in addressing disparities between socio-economically advantaged and disadvantaged students.

The TALIS-PISA link data provides an opportunity to identify teacher and school factors that may matter for equity in student performance since it offers an internationally comparable dataset combining information on key stakeholders. Yet, The TALIS-PISA link data also present important limitations. The link between teachers and students is established at the school and not at the class level. In other words, the data do not allow matching a teacher with her or his students; rather the data only permit matching a sample of teachers teaching 15-year-old students in a school with a sample of 15-year-old students of that same school. Therefore, information on teachers is averaged at the school level to be analysed together with students' outcomes. Given that teachers of the same school differ significantly in terms of their characteristics and practices, linking data by averaging teachers' variables at the school level triggers a considerable loss of information.[5] For the analyses of gender gaps in student achievement, student performances are also averaged by gender at the school level and not at the class level. Finally yet importantly, the cross-sectional design of the survey prevents the measurement of any teacher causal effects or distinguishing between short-term and long-term effects of teachers and schools on students' outcomes.

Do teacher and school factors matter equally for students with different academic performances?

Performance at school has long-term consequences for both students and society as a whole. For example, Hanushek and Woessmann show that a small increase in students' cognitive outcomes, as measured by PISA, can have a very large impact on a nation's economic growth. According to their estimates, a 25-point rise in the average PISA score of all OECD countries has the potential of generating a GDP gain of USD 115 trillion over the lifetime of a generation (OECD, 2010[9]). Concentrating efforts on raising the academic level of lower-performing students is an effective way to improve education systems' overall performance and equity since students from disadvantaged backgrounds are over-represented among low performers (OECD, 2018[5]; OECD, 2016[10]).

Students' academic performance at school is not associated with a single student, teacher or school characteristic; rather, it results from the combination of several factors. Chapter 2 of this report describes these factors extensively as well as their direct and indirect effects on student performance by focusing on the mean of student performance distribution, i.e. the average performing student.

However, by focusing on average performing students, these findings may miss important variations in the way teacher and school factors matter for students with different competence levels. To this end, this section applies quantile regression analysis to identify if the factors identified in Chapter 2 as key predictors of PISA test scores for average performing students differ at various points in the conditional distribution of students' performances (as measured in PISA 2018 across reading, mathematics and science). In contrast with standard linear regressions that focus on the average performing student, quantile regressions can provide a more detailed picture by allowing for analysis of the effect of teacher and school factors on student performance at different achievement levels (Box 4.1).

Box 4.1. Quantile regression

While standard linear regression estimates the conditional mean of the outcome variable, given a set of explanatory variables, quantile regression provides information about the association between the outcome variable and the explanatory variables at the different percentiles in the conditional distribution of the outcome variable (Koenker, 2017[11]; Koenker, 2005[12]; Koenker and Bassett, 1978[13]).

Quantile regression can be appealing in comparison to standard linear regression (Cameron and Trivedi, 2009[14]), since:

- It provides a richer characterisation of the relationship between the outcome variable and the explanatory variables by allowing the effects of the explanatory variables to vary over different quantiles of the conditional distribution
- It is more robust in terms of outliers and assumptions about the distribution of regression errors.

However, it is important to note that quantile regression estimates tend to be more precise at the centre of the distribution as compared to upper and lower quantiles. Thus, for those relationships where the variation in the effects of the explanatory variables over the different quantiles of the conditional distribution of the outcome variable is limited, quantile regression is less likely to find significant coefficients at the tails of a distribution than at the centre.

Note: For additional information on quantile regression, see Annex B.
Sources: Cameron, A. and P. Trivedi (2009[14]), *Microeconometrics Using Stata*; Koenker, R. (2017[11]), "Quantile regression: 40 years on", *Annual Review of Economics*, http://dx.doi.org/10.1146/annurev-economics-063016-103651; Koenker, R. (2005[12]), *Quantile Regression*, http://dx.doi.org/10.1017/CBO9780511754098; Koenker, R. and G. Bassett (1978[13]), "Regression quantiles", *Econometrica: The Journal of the Econometric Society*, http://dx.doi.org/10.2307/1913643.

Quantile regression analysis had not been applied on TALIS data or on TALIS-PISA link data before this report. Yet, various studies have estimated quantile regressions using PISA data, shedding light on more nuanced findings. A few studies focused on the varying relationship of school composition with student performance. For example, Rangvid (2007[15]) combined PISA 2000 data with administrative data from Denmark to estimate the relationships between school socio-economic composition and students' test scores. This study finds that low-achieving students benefit more from socially advantaged peers than high-achieving students do in reading performance but not in mathematics and science.[6] Also, Raitano and Vona (2013[16]) studied the interaction between school-tracking policies and peer effects. Exploiting PISA 2006 data, they concluded that peer heterogeneity in terms of ability reduces students' performances in the comprehensive system while it increases student performance in the early-tracking system. In addition, they also found evidence for stronger peer effects for low-ability students.

A few other studies looked at the varying relations between teacher or school practices and student performance. Using PISA 2015 data, Mostafa et al. (2018[17]) examined – among others – whether the relationship between science proficiency and teaching practices vary at different points in the distribution of students' science test scores. According to their findings, enquiry-based science and teaching practices and teacher-directed science instruction appear to matter as much for low performers as for top performers (i.e. students in the bottom decile and in the top decile of the science test score distribution) in all or almost all of the PISA countries and economies with available data. Drawing on PISA 2012 data, Teng (2020[18]) explored the effects of school climate on disparities in students' mathematics achievement in Shanghai (China) and concluded that school climate matters more in low and average-performing schools, and for underachievers and medium-level students. In an article based on PISA 2009 data for Thailand, Lounkaew (2013[19]) analysed the gap between students in urban and rural areas. Their results pointed to the importance of unobservable school characteristics beyond the student and school characteristics measured by PISA in explaining achievement gaps between these students. Decomposition exercises by level of achievement revealed the increasing role of these unobserved characteristics: the higher students' achievement levels are, the greater the role these characteristics play in explaining the achievement gap between urban and rural areas.

All in all, findings from past studies that applied quantile regressions on PISA data indicate that certain school contexts and practices may have differential effects on students, depending on their achievement level. This kind of analyses can reveal policy levers that would address the needs of targeted groups of students such as low achievers. They also identify teacher and school practices that are actually effective for all students, regardless of their ability level.

This section applies quantile regressions on the TALIS-PISA link dataset to investigate whether the effects of teacher and school factors differ along the distribution of students' test scores. Chapter 2 of this report concludes that factors such as teachers' classroom practices (in particular the share of class time spent on actual teaching and learning), classmates' characteristics (in particular the share of socio-economically disadvantaged students and the concentration of academically gifted students in the school), teachers' well-being and job satisfaction (in particular workload stress and satisfaction with the work environment), teachers' use of working hours (in particular the hours spent on marking and correcting student work) and school culture (in particular, stakeholders' involvement in school activities) are key predictors of students' PISA test scores in reading, mathematics and science. By estimating quantile regressions at the 25th and 75th percentiles in the conditional distribution of students' performances in reading, mathematics and science, this section examines whether the aforementioned factors also matter for low achievers and their high-achieving peers. It also explores whether there are practices and characteristics of teachers and schools that matter specifically for low-performing students or their high-performing peers.[7]

For ease of reading in the following sections, students at the 25th percentile (or equivalently at the bottom quartile) of the test score distribution are referred to as low achievers (or in equivalent terms as low performers, low-achieving students, etc.). Students at the 75th percentile (or at the top quartile) of the test score distribution are referred to as high achievers. In the following sections, the interpretation of the quantile regression results focuses on the significance and to a lesser extent on the size of the regression coefficients estimated at the bottom and the top quartiles of the conditional performance distribution. For example, a teacher factor will be deemed more important for low achievers than for high achievers if it is found to be significant (and generally greater in absolute value) for the bottom quartile but not for the top quartile. The reported analyses are likely to be conservative to the extent that an association between certain teacher and school factors and outcomes of low- or high-achieving students might not be significant due to less precision at the tails of the distribution (Box 4.1) even though these factors might truly contribute to the academic performance of low or high achievers. Differences in the coefficients between quartiles are neither reported nor commented on as in most cases they are not significant. Therefore, great caution must be taken while interpreting these findings.

Teachers' classroom practices

The teacher dimension of classroom practices is consistently highlighted in Chapter 2 as well as by education research as an important predictor of student performance (Hattie, 2009[20]; Le Donné, Fraser and Bousquet, 2016[21]; Muijs et al., 2014[22]). The more class time the average school teacher spends on actual teaching and learning, the better students tend to perform on average in all three subjects covered by PISA (i.e. reading, mathematics and science) in the Czech Republic, Turkey and on average across the countries and economies participating in the TALIS-PISA link (Tables 2.5, 2.9 and 2.13). In Denmark, there is a positive relationship for mathematics and science, but the association is not statistically significant for reading.

Results of quantile regression analysis (focusing on the 25th and 75th percentiles of the conditional distribution of students' performance in PISA tests) indicate that the share of class time spent on actual teaching and learning matters as much for low performers as for top performers on average across the countries and economies participating in the TALIS-PISA link and in Turkey (Figure 4.1). In the case of the Czech Republic, the positive relationship between teachers' time spent on actual teaching and learning and student performance seems to hold only for low-performing students. This suggests that only low performers benefit from additional teaching and learning time in the Czech Republic.

Research on the relationship between learning time and student achievement offers mixed evidence. Patall, Cooper and Allen (2010[23]) conducted a comprehensive search of the literature on this topic and found 15 studies conducted between 1985 and 2009. The literature revealed, first, that survey designs are generally weak for making causal inferences. The TALIS-PISA link design is no exception to this. That said, findings from the literature suggest that extending school time can be an effective way to support student learning, particularly for students most at risk of school failure and when considerations are made for how time is used. Moreover, the strongest research designs produced the most consistent positive results. Here, it must be kept in mind that what the TALIS-PISA link measures is the share of typical lessons spent on actual learning and teaching and not the actual amount of time spent on learning by students. Findings suggest that such quality class time is beneficial to all students, regardless of their ability level. This suggests that teachers should spend as much as possible of their class time on teaching and learning rather than on administrative tasks and classroom management. The exception found in the Czech Republic is consistent with past research finding that additional learning time is especially important for slower learners and low-performing students (Gromada and Shewbridge, 2016[24]).

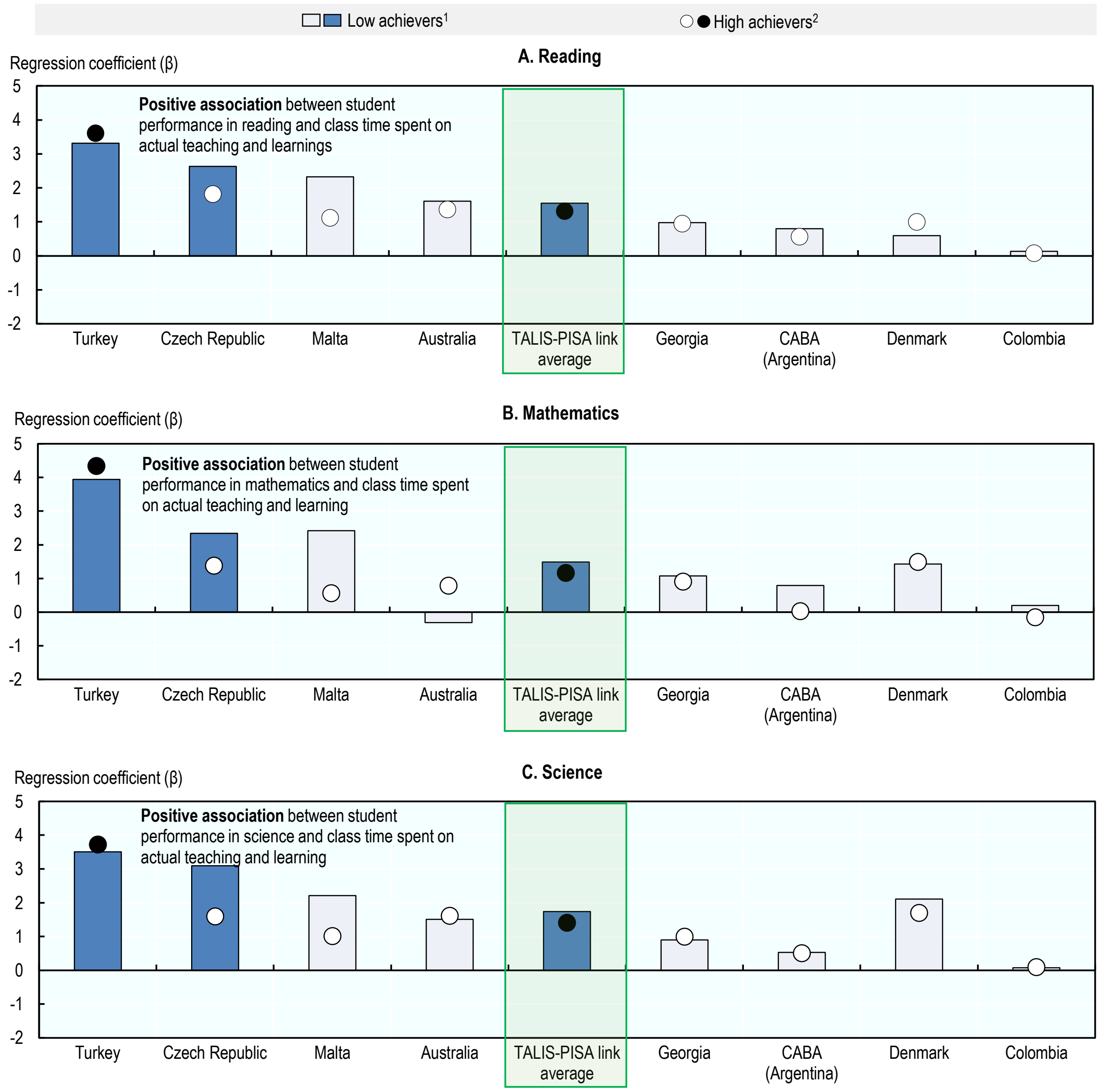
In the case of Denmark, however, the association between the share of class time spent on actual teaching and learning, and student performance is no longer significant when moving away from average performing students and focusing on low and high achievers (Figure 4.1). This suggests that teachers, deliberately or not, tailor their teaching practice and use of time to address the needs of average students. However, this finding may be due to a statistical artefact of quantile regression analysis as estimates tend to be more precise at the centre of the distribution as compared to upper and lower quantiles (Box 4.1).

Teachers' well-being and job satisfaction

In line with previous research findings, which indicate that teachers' satisfaction with their job can have an indirect positive effect on student achievement through teachers' beliefs, attitudes and practices as well as school culture (Ainley and Carstens, 2018[25]), teachers' well-being and job satisfaction is a dimension that is found to matter for student performance in Chapter 2. The more teachers are satisfied with the work environment on average at the school, the better students tend to perform in all three subjects covered by PISA in Australia, the Czech Republic, Georgia, Turkey and on average across participating countries and economies (Tables 2.25, 2.29 and 2.33).

Figure 4.1. Conditional relationship between class time spent on teaching and student achievement

Change in PISA score associated with average class time spent on actual teaching and learning at the school, conditional on student PISA score, by subject



1. Low-achieving students are those at the bottom quartile (i.e. 25th percentile) of the conditional distribution of PISA score.
2. High-achieving students are those at the top quartile (i.e. 75th percentile) of the conditional distribution of PISA score.

Notes: Results of quantile regression based on responses of 15-year-old students and teachers. Controlling for the following classroom practices of teachers: teachers' autonomy over planning and teaching, perceived disciplinary climate, use of practices related to clarity of instruction, use of practices related to cognitive activation, use of assessment practices, such as administering own assessment, providing written feedback on student work in addition to marking, letting students evaluate their own progress and observing students when working on particular tasks and providing immediate feedback; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. Teacher variables are averaged for all teachers within the school. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

Countries and economies are ranked in descending order of the change in low achievers' PISA score in reading associated with the average class time spent on actual teaching and learning at the school.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6.

StatLink https://doi.org/10.1787/888934224030

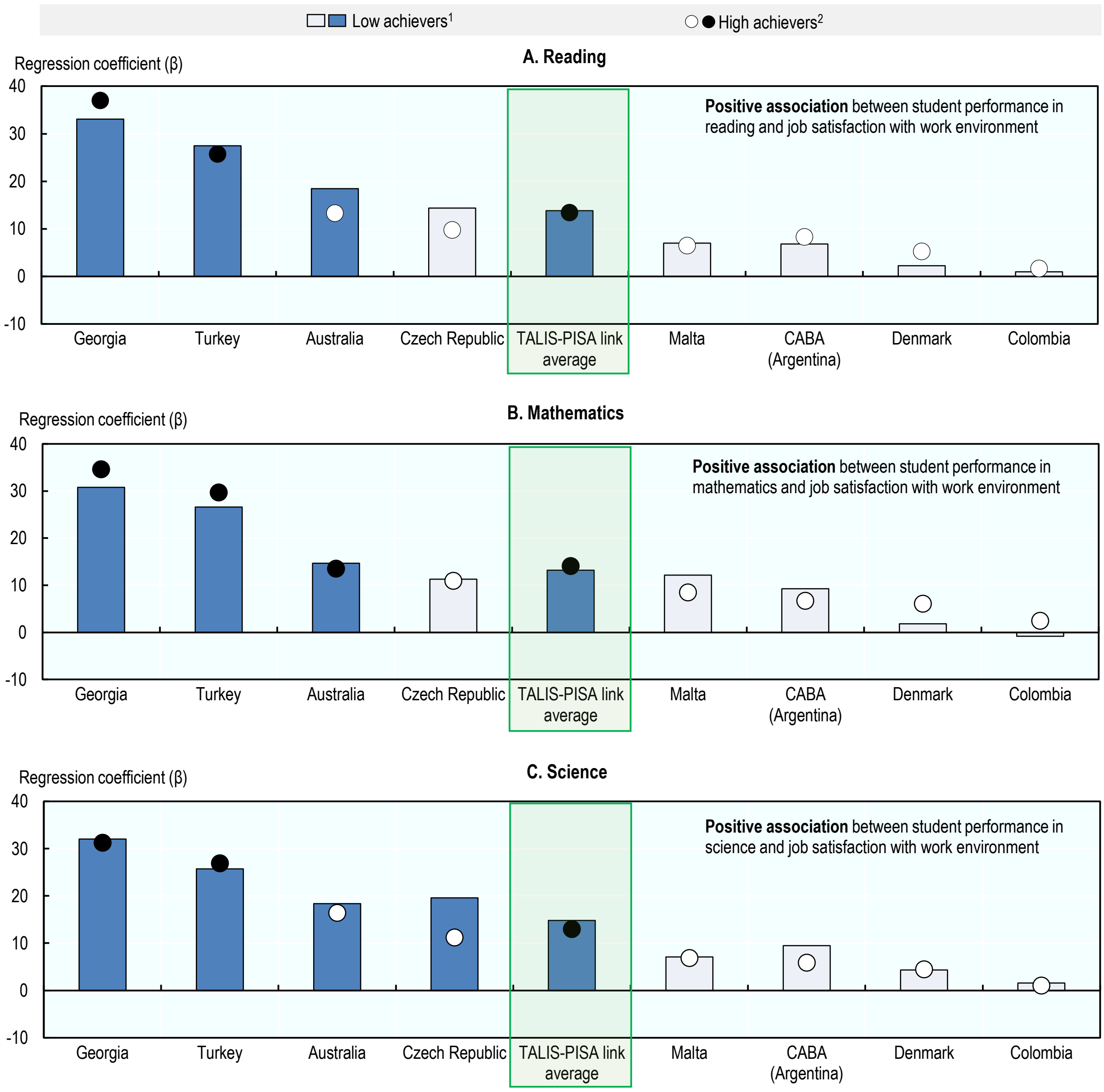
Quantile regression analysis indicates that, regardless of the subject, teachers' satisfaction with the work environment is positively associated with student performance for both low- and high-performing students on average across participating countries and economies, and in Georgia and Turkey (Figure 4.2). The same result holds in the case of Australia, but only for mathematics. In Australia, the relationship between teachers' satisfaction with work environment and student performance seems to matter more for low achievers when it comes to reading and science. In the case of the Czech Republic, the association is significant only when focusing on low performers in science. The results found for Australia and the Czech Republic may signal that teachers' job satisfaction matters more for low performers than for their high-achieving peers. Teachers who are satisfied with their work environment tend to report higher self-efficacy (OECD, 2020[26]; OECD, 2014[27]), while job satisfaction also plays an important role in teachers' attitudes and efforts (Ainley and Carstens, 2018[25]). Thus, teachers' satisfaction with their work environment may be especially beneficial for low performers through teachers' increased self-efficacy and commitment.

Besides teachers' satisfaction with the work environment, the extent to which workload is an important source of stress is also an important predictor of student achievement. In Colombia, the Czech Republic, Denmark and on average across participating countries and economies, students who attend schools where teachers, on average, report workload (including lesson preparation, number of lessons to teach, marking, administrative work and extra duties due to absent teachers) as an important source of stress, tend to perform better in all three subjects covered by PISA (Tables 2.25, 2.29 and 2.33). This may signal teachers' commitment and dedication to their work as well as point to a reverse causal relationship as highly competitive school environments (attended by higher- performing students) can lead to workload being an important source of stress for teachers.

Quantile regressions results show that the positive association between workload being an important stress for teachers and student performance holds for both low and top performers in the Czech Republic for all three subjects covered by PISA, only for mathematics in Colombia, only for reading in Denmark and only for science on average across participating countries and economies (Figure 4.3). In Colombia, workload-induced stress seems to matter more for high-achieving students in reading and science. This may hint at the presence of highly competitive school environments that can lead to additional workload-induced stress for teachers. On the contrary, in the case of Denmark, the relationship seems to be more prominent for low performers in mathematics and science, which may signal teachers' commitment and dedication to their work in general. More concretely, it may hint at the extra effort teachers make in helping low-achieving students, which, in turn, can lead to additional workload-induced stress. Similarly, workload-induced stress seems to be more closely related to student achievement for low-performing students in reading and mathematics on average across participating countries and economies.

Figure 4.2. Conditional relationship between teachers' satisfaction with work environment and student achievement

Change in PISA score associated with the average job satisfaction with work environment at the school, conditional on student PISA score, by subject

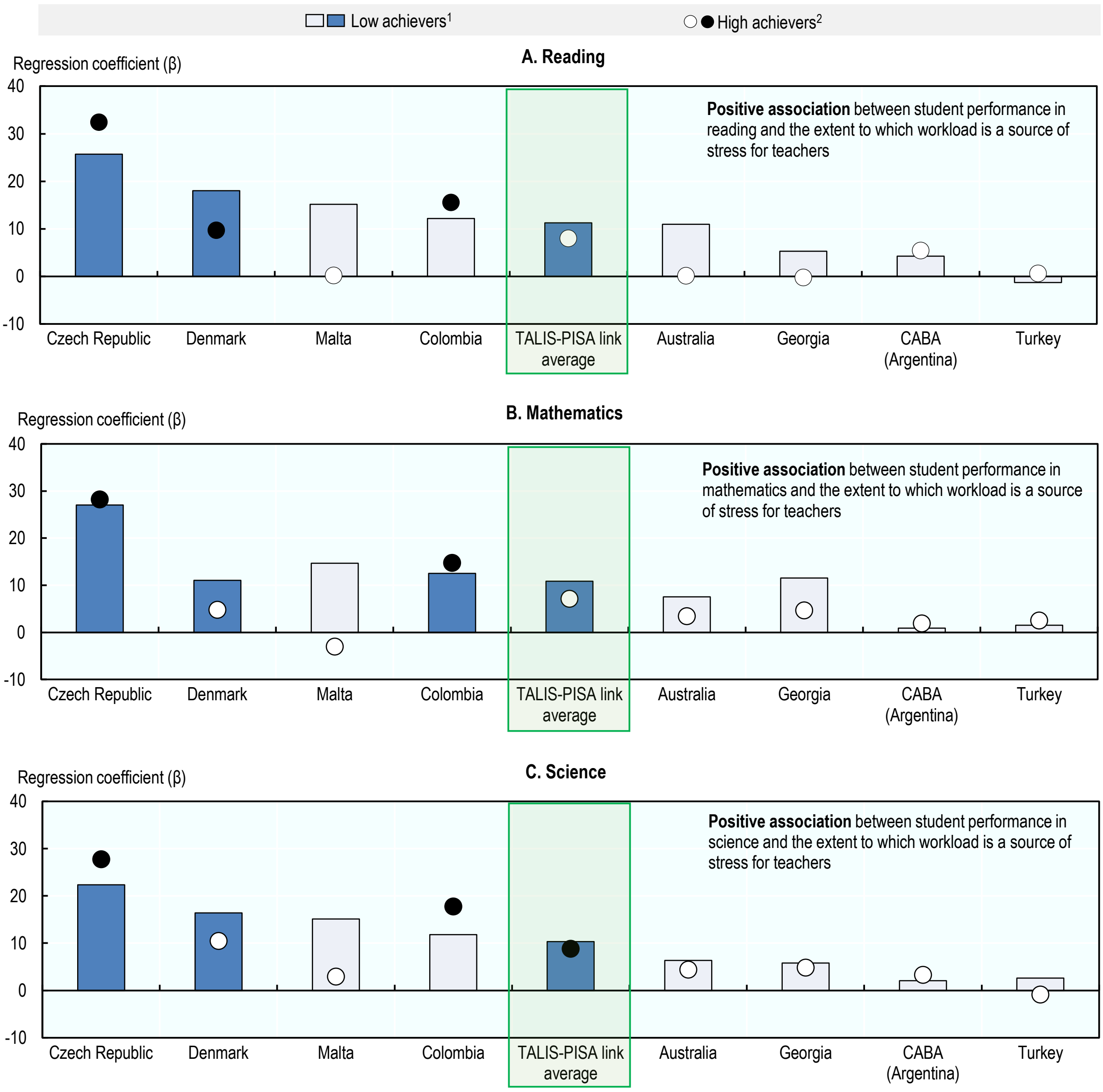


1. Low-achieving students are those at the bottom quartile (i.e. 25th percentile) of the conditional distribution of PISA score.
2. High-achieving students are those at the top quartile (i.e. 75th percentile) of the conditional distribution of PISA score.
Notes: Results of quantile regression based on responses of 15-year-old students and teachers. Controlling for the following aspects of well-being and job satisfaction: workplace well-being and stress, workload stress, job satisfaction with profession, teachers' satisfaction with the salary, teachers' satisfaction with the terms of the teaching contract apart from salary (e.g. benefits, work schedule), teachers' views of the way different stakeholders value the profession; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. Teacher variables are averaged for all teachers within the school. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in low achievers' PISA score in reading associated with the average job satisfaction with work environment.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.7, 4.8, 4.9, 4.10, 4.11 and 4.12.

StatLink https://doi.org/10.1787/888934224049

Figure 4.3. Conditional relationship between teachers' workload-induced stress and student achievement

Change in PISA score associated with the average extent to which workload is a source of stress for teachers at the school, conditional on student PISA score, by subject



1. Low-achieving students are those at the bottom quartile (i.e. 25th percentile) of the conditional distribution of PISA score.
2. High-achieving students are those at the top quartile (i.e. 75th percentile) of the conditional distribution of PISA score.

Notes: Results of quantile regression based on responses of 15-year-old students and teachers. Controlling for the following aspects of well-being and job satisfaction: workplace well-being and stress, job satisfaction with the work environment, job satisfaction with profession, teachers' satisfaction with the salary, teachers' satisfaction with the terms of the teaching contract apart from salary (e.g. benefits, work schedule), teachers' views of the way different stakeholders value the profession; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. Teacher variables are averaged for all teachers within the school. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

Countries and economies are ranked in descending order of the change in low achievers' PISA score in reading associated with the average extent to which workload is a source of stress for teachers at the school.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.7, 4.8, 4.9, 4.10, 4.11 and 4.12.

StatLink https://doi.org/10.1787/888934224068

Teachers' use of working time

The way teachers balance their time among the often competing tasks is important for the quality of the teaching and student learning (OECD, 2019[28]). While research highlights the positive implications of formative assessment in the form of constructive and immediate feedback for teaching and learning (Muijs and Reynolds, 2001[29]), summative assessment of students' work in the form of tests and exams also provides feedback to students about their learning progress (Ainley and Carstens, 2018[25]). In addition, teachers' time spent on summative assessment can be indicative of teachers' engagement in formative assessment. According to the results of multiple linear regressions in Chapter 2, the more working hours teachers spend on marking and correcting, the better students tend to perform on average at the school in all three subjects covered by PISA in Australia, Colombia, the Czech Republic, Georgia and on average across participating countries and economies (Tables 2.37, 2.41 and 2.45).

The positive relationship between teachers' working time spent on marking and correcting student work, and student performance observed for the average performing student seems to matter as much for low performers as for top performers, irrespective of the subject, in Colombia, the Czech Republic, Georgia and on average across participating countries and economies (Tables 4.13, 4.14, 4.15, 4.16, 4.17 and 4.18). This is also true for Australia, but only for reading. In Australia, teachers' working time spent on marking and correcting seems to matter especially for low achievers in mathematics and science, and not significantly for their high-achieving peers.

Classmates' characteristics

Chapter 2 of this report suggests the presence of peer effects. TALIS-PISA link data show that as the average concentration of students from socio-economically disadvantaged homes in the classrooms increases, the worse students tend to perform academically in several countries and economies participating in the TALIS-PISA link, including Australia, Colombia and Georgia (Tables 2.49, 2.51 and 2.53). This relationship also becomes significant for the TALIS-PISA link average for all three subjects as the focus shifts exclusively to subject teachers (Tables 2.50, 2.52 and 2.54). Indeed, a student's performance can be negatively affected if surrounded by classmates with limited social, economic and cultural resources. This may also reflect an overall concentration of important disadvantages in the student's local community. As socio-economic disadvantage often translates into lower achievement, being surrounded by classmates from socio-economically disadvantaged homes tends to go hand-in-hand with having academically struggling classmates. Past research shows that struggling classmates can be detrimental to student performance due to reduced teaching time and altered teaching strategies (OECD, 2019[7]; Sacerdote, 2011[30]). Yet, former research findings also suggest that sorting students by their ability may widen disparities in performance since high-performing students tend to be less affected than their low-achieving peers by the composition of their classes (Burke and Sass, 2013[31]; Lavy, Silva and Weinhardt, 2012[32]; Sacerdote, 2011[30]).

Quantile regression results indicate that the average share of students from socially disadvantaged homes in the classroom is negatively related to performance of both low- and high-performing students in Colombia (Figure 4.4). In Australia, the concentration of disadvantage appears to matter specifically for low-achieving students in mathematics performance. This finding is in line with the research consensus that suggests that low achievers may be more sensitive to the composition of their classes compared to their high-performing peers (Burke and Sass, 2013[31]; Lavy, Silva and Weinhardt, 2012[32]; Sacerdote, 2011[30]). By contrast, classmates' characteristics seem to matter only for high-achieving students in Georgia, irrespective of the subject, and in Turkey in the case of science. In the remaining countries and economies participating in the TALIS-PISA link, there is no significant relationship found between the share of socio-economically disadvantaged students and student performance, neither for low achievers nor for their high-achieving peers. This may be an artefact of focusing on the lower and upper quartiles in the distribution of student performance as opposed to the average performing student since quantile regression analysis is less likely to find significant results for the extremes of a distribution than its centre (Box 4.1). Moreover, as stated by van Ewijk and Sleegers (2011[33]), the effect of peers' socio-economic status on student achievement is highly dependent on how the socio-economic composition of a school or a classroom is defined.

The share of academically gifted students in the classroom also matters for student achievement. The greater the average concentration of academically gifted students in the classroom, the better students tend to perform in all subjects covered by PISA in Australia, CABA (Argentina), the Czech Republic, Turkey and on average across participating countries and economies (Tables 2.49, 2.51 and 2.53). The only country where the association between the average concentration of academically gifted students in the classrooms and student achievement is not significant for any of the subject domains is Colombia. These results signal the presence of academic segregation but they can also point to the potential presence of peer effects. A student's performance can be positively affected by classmates with higher innate ability through an increase in motivation, competition and career aspirations (OECD, 2020[26]; Sacerdote, 2011[30]).

Quantile regression results show that the average concentration of academically gifted students in the classrooms matters, irrespective of the subject, as much for low performers as top performers in CABA (Argentina), the Czech Republic, Turkey and on average across countries and economies participating in the TALIS-PISA link (Tables 4.19, 4.20, 4.21, 4.22, 4.23 and 4.24). This is also true for Australia, but only for mathematics and science. In Denmark, the concentration of students with innate ability appears to matter specifically for high-achieving students in mathematics, and is insignificant in reading and science.

School culture

The dimension of school culture is highlighted by former research as well as findings in Chapter 2 as being closely related to student performance (Ainley and Carstens, 2018[25]). Students who attend schools where stakeholders (i.e. parents and community) are involved in school-related activities tend to perform better in the PISA test across all subjects in Australia, Colombia, Denmark and on average across participating countries and economies (Tables 2.55, 2.59 and 2.63). These findings are in line with past research that shows a positive association between parental and community involvement, and student achievement (Wang and Degol, 2016[34]; Wilder, 2014[35]).

Quantile regression results indicate that, on average across participating countries and economies and in Denmark, the positive association between stakeholder involvement in school-related activities and student performance holds both for low and top performers in all three subjects covered by PISA (Figure 4.5). In Colombia, the same relationship is observed for reading and mathematics. However, when it comes to performance in science, stakeholder involvement seems to matter specifically for high achievers. In Australia, the positive association only holds for reading. In the case of mathematics, high-achieving students are the ones who seem to benefit from the involvement of parents and the local community while the relationship between stakeholder involvement and student performance in science is significant neither for low-achieving students nor for their high-achieving peers.

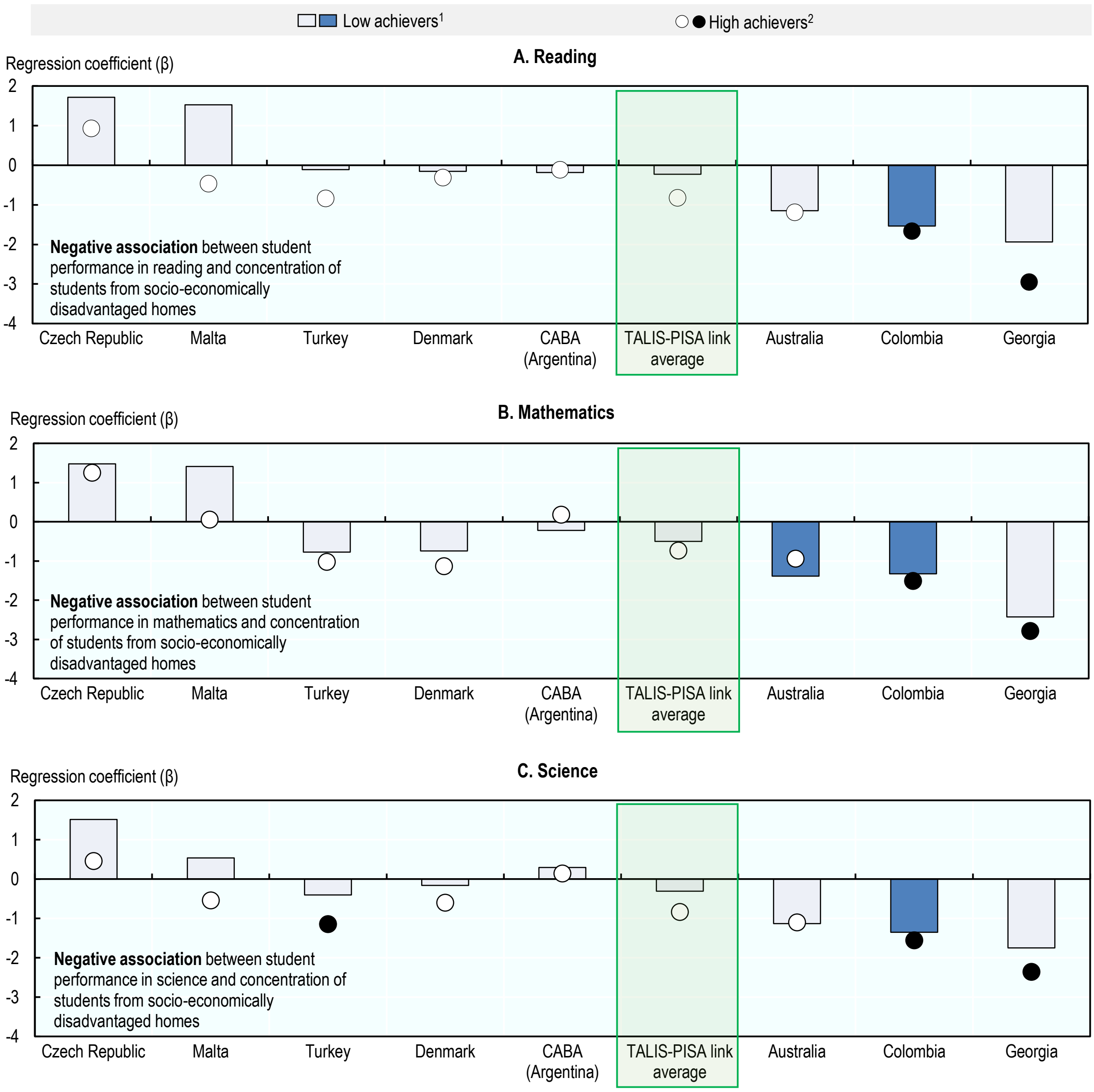
Summary

Results presented above show that teacher and school factors that matter for average students' performances in reading, mathematics and sciences also tend to matter for both low and high-achieving students on average across the countries and economies participating in the TALIS-PISA link. Teacher factors include the share of class time that teachers report spending on actual teaching and learning, the working hours teachers report devoting to correcting and marking their student work, and teachers' satisfaction with the work environment. In addition, the average concentration of academically gifted students in the classroom as well as the degree of parents' and the community's involvement in school-related activities also matter for all students, regardless of their performance level. This suggests that all these practices and characteristics could be leveraged to equally support student academic growth, regardless of their initial competence level.

One teacher factor is found to matter specifically for low-achieving students: workload-induced stress. The share of teachers reporting workload as an important source of stress at the school is positively associated with better performances for low-achieving students in two out of the three subject domains (reading and mathematics). This may signal teachers' commitment and dedication to their work in general. More concretely, it may hint at the extra effort teachers put in helping low-achieving students, which, in turn, can lead to additional workload-induced stress. While stress can be a potent force and reflect a feeling of commitment and dedication, it can also develop into burnout. Thus, workload-induced stress may signal the need for support for committed and dedicated teachers who feel overwhelmed with the workload.

Figure 4.4. Conditional relationship between concentration of students from socio-economically disadvantaged homes and student achievement

Change in PISA score associated with the average concentration of students from socio-economically disadvantaged homes at the school, conditional on student PISA score, by subject



1. Low-achieving students are those at the bottom quartile (i.e. 25th percentile) of the conditional distribution of PISA score.
2. High-achieving students are those at the top quartile (i.e. 75th percentile) of the conditional distribution of PISA score.

Notes: Results of quantile regression based on responses of 15-year-old students and teachers. Controlling for the following classroom characteristics: class size, share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, academically gifted students, students who are immigrants or with a migrant background and students who are refugees; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. Teacher variables are averaged for all teachers within the school. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

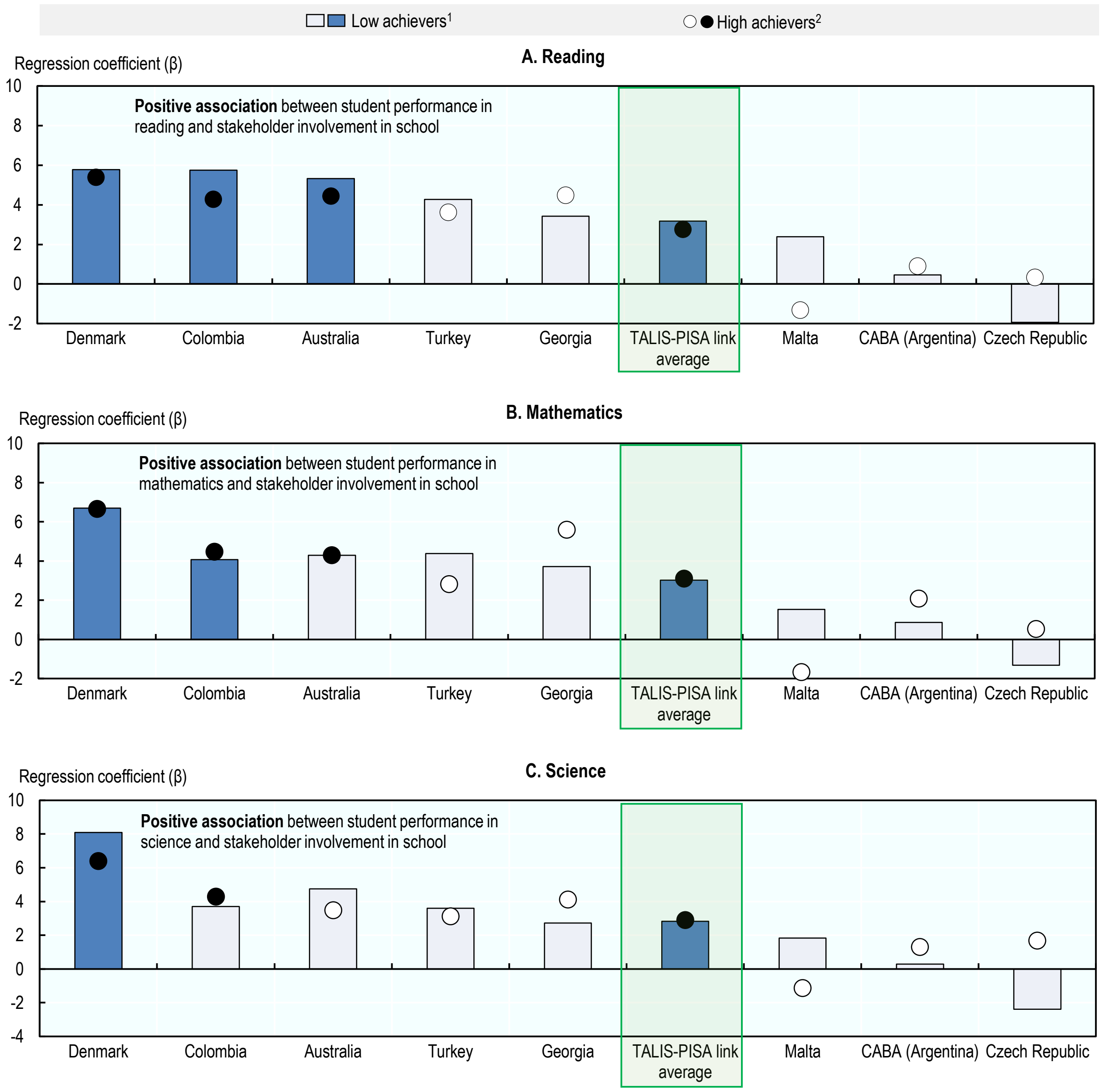
Countries and economies are ranked in descending order of the change in low achievers' PISA score in reading associated with the average concentration of students from socio-economically disadvantaged homes at the school.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.19, 4.20, 4.21, 4.22, 4.23 and 4.24.

StatLink https://doi.org/10.1787/888934224087

Figure 4.5. Conditional relationship between stakeholder involvement in school and student achievement

Change in PISA score associated with stakeholder (i.e. parents and local community) involvement in school, conditional on student PISA score, by subject



1. Low-achieving students are those at the bottom quartile (i.e. 25th percentile) of the conditional distribution of PISA score.
2. High-achieving students are those at the top quartile (i.e. 75th percentile) of the conditional distribution of PISA score.
Notes: Results of quantile regression based on responses of 15-year-old students, teachers and principals. Controlling for the following aspects of school culture: collaborative school culture, teacher-student relations and teachers' actions towards achieving academic excellence; and for the following student characteristics: gender, immigrant background and index of economic, social and cultural status. Teacher variables are averaged for all teachers within the school. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in low achievers' PISA score in reading associated with stakeholder involvement in school.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.25, 4.26, 4.27, 4.28, 4.29 and 4.30.

StatLink https://doi.org/10.1787/888934224106

How can schools mitigate gender disparities in student performance?

Gender disparities often have long-term consequences on students' fulfilment of their personal and professional potential (OECD, 2019[7]; UNESCO, 2018[36]). Mitigating these disparities in students' outcomes is high in the education policy agenda. PISA has consistently found that girls outperform boys in reading and, although to a lesser extent, that boys outscore girls in mathematics (OECD, 2019[7]). Boys tend to be overrepresented among students who lag behind and lack basic proficiency in reading that is necessary to meet the challenges of today's knowledge societies (OECD, 2019[7]). On the other hand, girls are usually not among top performers in science and mathematics, which, in turn, partly explains the underrepresentation of women in careers in science, technology, engineering and mathematics (STEM) fields (OECD, 2019[7]).

There is a growing body of evidence that concludes that school composition, school learning environment and some teacher practices are associated with gender gaps in performance. In particular, it seems that the school learning environment is related to the underperformance of boys in reading, and that girls in same-sex schools may perform better in mathematics and be more willing to take risks in their school work (OECD, 2015[8]). The social and economic resources that schools and families can provide for children's cognitive development appear to affect boys and girls differently. Boys from disadvantaged families have lower achievement scores and are less likely to complete high school than girls from a similar socio-economic background (Autor et al., 2016[37]; Brenøe and Lundberg, 2018[38]; DiPrete and Buchmann, 2013[39]). When it comes to teaching practices, PISA results suggest that teachers' use of cognitive-activation strategies in mathematics courses as well as teachers' use of strategies aimed at stimulating their students' enjoyment of reading are more effective among girls than among boys (OECD, 2015[8]).

Socialisation may also play a role in widening or bridging the gap in performances at school between girls and boys. Disparities in learning outcomes may be the result of how parents and teachers interact with boys and girls, which can be different. For instance, teachers may hold certain beliefs about boys' and girls' interests and abilities that may bias their own evaluations of student performance, which, in turn, may strengthen, or attenuate, gender gaps in achievement (Hadjar et al., 2014[40]; Robinson-Cimpian et al., 2014[41]). There is also some evidence that gender gaps in mathematics are smaller in cultures with weaker gender stereotypes (Else-Quest, Hyde and Linn, 2010[42]; Guiso et al., 2008[43]; Nollenberger, Rodríguez-Planas and Sevilla, 2016[44]; Nosek, Banaji and Greenwald, 2002[45]).

Research literature also suggests that teachers' gender can influence student performance. One broad set of explanations involves what is sometimes referred to as "passive" teacher effects. These effects are triggered by a teacher's gender identity, not by explicit teacher behaviour (Dee, 2005[46]). The most common example is the role-model effect, which operates when the presence of a teacher raises the interest and performance of students of a given gender in comparison to the other gender. For example, research was conducted on students in the US Air Force Academy, who were randomly assigned to a female professor in mandatory introductory mathematics and science courses (Carrell, Page and West, 2010[47]). The results suggest that although having a female professor has little impact on male students, it has a powerful effect on female students' performance in mathematics and science, and high-performing female students' decision to continue studying these subjects in the future. Another large-scale experiment conducted in the Paris area in France showed that a brief classroom intervention with female engineers can significantly reduce gender gaps in performance and the prevalence of stereotypical views on jobs in science (Breda et al., 2020[48]).

A second set of explanations points to "active" teacher effects, that is, biases – whether deliberate or unintended – in their prior expectations of and interactions with students who are boys or girls (Dee, 2005[46]). For instance, a study that randomly assigned teachers and students to classes in primary schools in Tel-Aviv, Israel, found that a teacher with a greater bias[8] in favour of girls (boys) has positive effects on girls' (boys') performances (Lavy and Sand, 2018[49]). More specifically, even while controlling for students' behaviour and work ethics, as well as for past and current test scores, research suggests that teachers tend to underrate girls' mathematics proficiency. This, in turn, accounts for a substantial share of the mathematics achievement gap between equally performing and behaving boys and girls (Cimpian et al., 2016[50]; Robinson-Cimpian et al., 2014[41]).

In addition, some observational studies indicate that peer influence operates differently among boys and among girls, and that boys are exposed to greater peer pressure to conform to gender identities than girls (OECD, 2015[8]). For boys, gender identity is marked by a relative lack of interest in school in general, and in reading in particular; on the other hand, gender identity appears to have a substantial negative impact on girls' interest in mathematics (OECD, 2015[8]). Gender stereotypes seem to affect students' self-confidence and, through this channel, gaps in performances (Carlana, 2019[51]). Research also suggests that girls' low self-belief in their abilities remains, even if they perform equally well, or better, than boys (Parker, Van Zanden and Parker, 2018[52]).

Drawing on TALIS-PISA link data, this section explores the teacher and school factors that could play a role in mitigating within-school disparities in performance between girls and boys. As in Chapters 2 and 3, this section uses a supervised statistical learning method, lasso, to investigate the relationship between teacher and school factors, and within-school disparities in student outcomes (see Box 2.1 for more information). It builds on the same list of potential predictors (almost 150 predictors across 18 teacher and school dimensions) to explain the school-level differences between girls and boys. Moreover, lasso regression results are complemented with a country-level analysis that aims to identify the relative importance of each teacher and school dimension in relation to within-school disparities. To this end, multiple regression analysis featuring teacher and school dimension separately is applied to establish the teacher and school dimensions that explain the highest shares of variance in within-school disparities.

School-level differences in performance between girls and boys are regressed on indicators of each teacher and school dimension (taken separately) that is either flagged by lasso regressions as an important predictor or explains a non-negligible part of variance in within-school disparities in student performance (i.e. 20% or above on average across countries and economies). In the context of this chapter, multiple linear regressions are estimated on one dimension at a time and provide insights into how the value of within-school differences in student performance changes when any one of the independent variables within a dimension varies while all other independent variables included in the model are held constant. In comparison to lasso regressions, multiple linear regressions provide the confidence intervals of the coefficient estimates, which, in turn, allow for drawing inferences about the overall population. Moreover, they also lead to more accurate coefficient estimates through the introduction of final and balanced repeated replicate weights and the use of plausible values of student performance. In contrast to the lasso regressions, which are based on the overall population of students, teachers and principals surveyed within the TALIS-PISA link, multiple linear regressions are applied at the country level and, as a result, allow for establishing cross-country patterns.

In order to examine the teacher and school factors that matter for within-school disparities in performance by students' gender, the analyses within this section are conducted at the school level. To this end, TALIS and PISA data are linked by merging student data averaged at the school level, which are collected by PISA, with teacher data averaged at the school level and principal data, which are collected by TALIS. As student data are averaged at the school level, controls only include school-level average student characteristics but not the average school-level classroom composition. Moreover, as the analyses are conducted at the school level rather than at the student level, as in Chapters 2 and 3 of this report, sample sizes decrease considerably. This is a limitation in so far as the identification of statistically significant relationships between teacher and school factors, and within-school disparities in student outcomes is more challenging compared to student-level analysis.

The next section investigates the teacher and school factors that are significantly related to within-school disparities in performance between girls and boys. It first focuses on gender disparities by discussing the direction and size of school-level gender gaps. School-level disparities in outcomes between girls and boys are defined as the difference between the average school-level PISA score for girls minus the average school-level PISA score for boys. Therefore, differences are positive when they are in favour of girls and negative when they are in favour of boys.

The chapter then explores the relationship between the many teacher and school dimensions measured in TALIS and within-school disparities in student outcomes measured in PISA. It then examines some specific hypotheses related to policies that could play a role in reducing gender disparities in student achievement. Namely, this section looks at whether male teachers could help close the gap in reading and whether female teachers could help close the gap – especially when it comes to mathematics performance – through distinct teaching practices or by acting as role models. To harness the richness of the TALIS principal questionnaire, which includes specific questions related to gender equity, the section also explores the association between certain school policies and practices as well as teachers' attitudes to gender equity and student performance.

Differences in performance between girls and boys within schools

The inspection of TALIS-PISA link data shows that although the direction of the school-level gender gap varies across schools, even within the same country, there is a general pattern for each subject. In addition, this general pattern is mostly aligned with the country-level gaps in student performance as documented by PISA 2018 results. Notably, while girls tend to perform better than boys in reading, boys outperform, albeit to a lesser extent, girls in mathematics (OECD, 2019[7]). Gender disparities in science are negligible (OECD, 2019[7]).

In the case of student outcomes in reading, female students tend to significantly outperform their male peers in more than one-third of the schools (36%) on average across TALIS-PISA link countries and economies (Figure 4.6). However, there are 18% of schools on average across participating countries with same-gender students where within-school gender gaps cannot be estimated.[9] In CABA (Argentina), Denmark, Georgia and Turkey, when focusing solely on significant differences, girls perform better in reading than boys in more than 40% of the schools. On the other end of the spectrum, in Colombia, the share of schools with a significant gap in favour of girls is 30%. Moreover, in Malta only 7% of schools present a significant gap. Yet, Malta also stands out among TALIS-PISA link countries and economies in that around 60% of its schools are single-gender schools. Apart from CABA (Argentina) (1%) and the Czech Republic (8%), all other countries and economies have between 10% and 16% of schools with same-gender students. Subsequently, the main challenge of small sample sizes in relation to school-level analysis is an especially important concern in the case of Malta.[10] As the analyses of the following sections all consist of school-level analyses by country, they will not include results for Malta, given the very small number of available cases.

The gender pattern for mathematics and science is opposite to that of reading. On average across TALIS-PISA link countries and economies, boys significantly outperform girls in mathematics in 20% of the schools (Figure 4.6). In CABA (Argentina), Colombia and the Czech Republic, when focusing solely on significant differences, boys outperform girls in mathematics in more than 25% of the schools. In science, the difference between girls and boys is less pronounced. Boys tend to significantly outperform girls in science in 16% of the schools, while the opposite pattern is observed in 11% of the schools, on average across participating countries and economies. Boys perform better than girls in science in 20% of the schools or more in the aforementioned countries and economies (CABA [Argentina], Colombia and the Czech Republic).

Average within-school gaps in performance between female and male students for each country and economy participating in the TALIS-PISA link corroborates the finding that gender disparities are most important in reading. The average within-school difference in reading between girls and boys is 19 score points in favour of girls, while the gap in favour of boys is 13 score points for mathematics and 6 score points for science, on average across participating countries and economies (Table 4.34). Average gender disparities in reading performance within schools is significant in six countries and economies, including Australia, CABA (Argentina), the Czech Republic, Denmark, Georgia and Turkey. In contrast, in mathematics, boys outperform girls on average within schools in four countries and economies – Australia, CABA (Argentina), Colombia and the Czech Republic. In the case of science, there are only two countries, Colombia and the Czech Republic, where a significant difference between girls and boys can be observed on average within schools.

However, TALIS-PISA link data suggest that there is room for schools to close gender gaps. The difference in reading performance between girls and boys is not statistically significant in 37% of schools on average in TALIS-PISA link countries and economies (Figure 4.6). Depending on the country/economy, this share varies between 27% (Malta) and 47% (CABA [Argentina]). Gender gaps in mathematics and science are not statistically significant in 55% of schools, on average across participating countries and economies. Of note, boys are not only able to perform equally well as (or not significantly differently than) girls in some schools but significantly outperform girls in reading in 9% of the schools. This share varies between 3% of the schools in Turkey and 14% of schools in Colombia and the Czech Republic (Figure 4.6). Similarly, girls are also able to perform better than boys in mathematics and even more so in science. Girls significantly outperform boys in mathematics in 7% of the schools on average across TALIS-PISA link countries and economies. In Australia and Turkey, this share increases to 11% while, in Colombia, there are almost no schools where girls seem to perform significantly better in mathematics than boys. The pattern is more balanced for science with more schools in which girls significantly outperform boys in science (11%; TALIS-PISA link average). There is even one country – Georgia – where the share of schools where girls outperform boys in science (21%) is greater than the share of schools where boys perform better (9%).

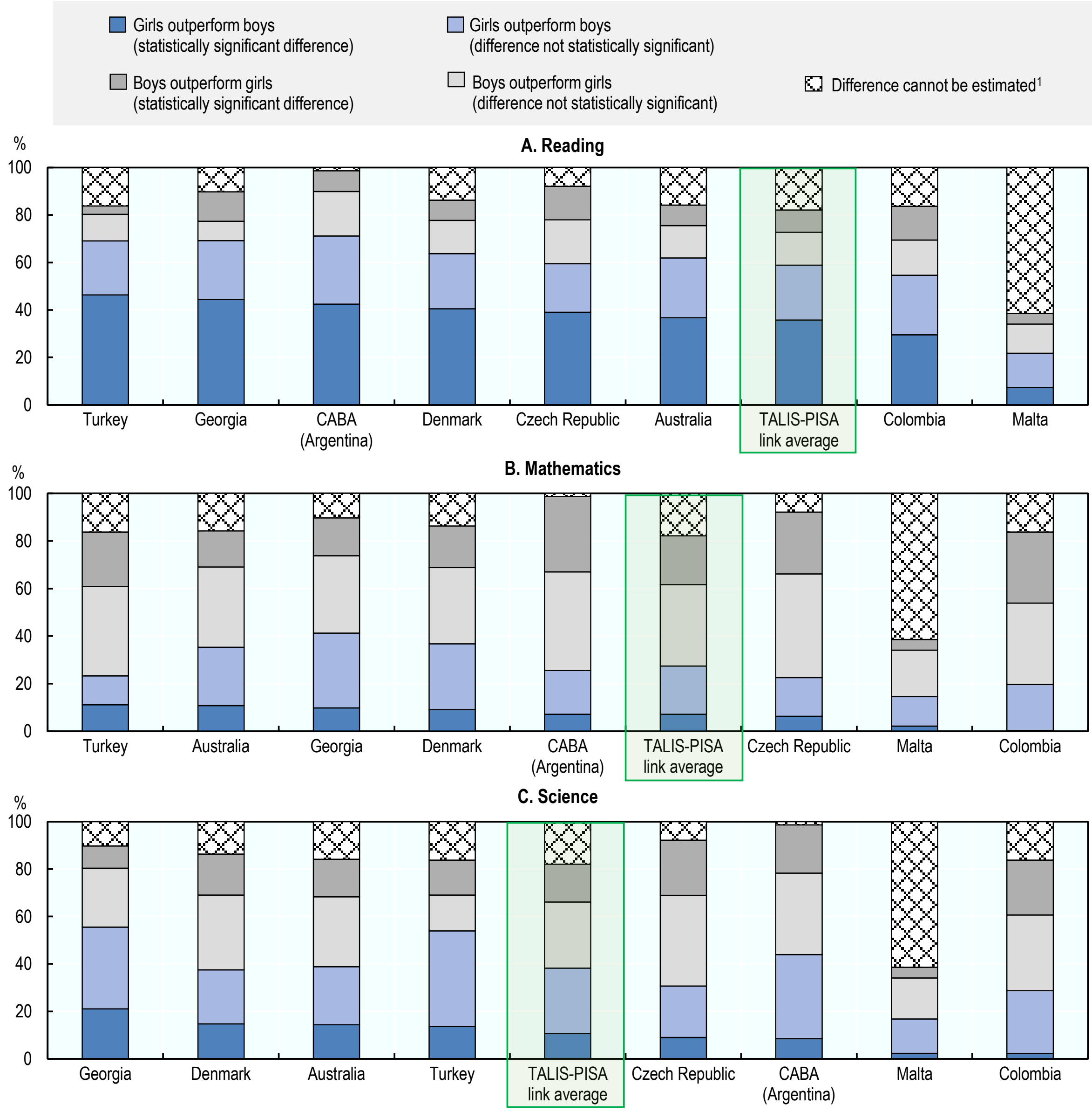
Teacher and school factors that matter for gender equity in achievement within schools

In each round of PISA, one out of the three domains of competence – reading, mathematics and science – is tested in detail, taking up nearly half of the total testing time. In 2018, reading was the focus. As reading literacy is measured more in-depth compared to mathematics and science within the latest TALIS-PISA link data, it is considered the most suitable subject for analyses applied on small samples. Therefore, this section focuses on student performance in reading. Due to low sample sizes, this section considers all teachers, not only reading teachers, when averaging teacher data at the school level. Focusing solely on reading teachers would lead to further loss of observations as that approach would require the exclusion of schools without any reading teachers sampled from the analysis.

Lasso regressions provide a data-driven approach to identifying which of the almost 150 potential predictors of the 18 teacher and school dimensions are significantly related to gender disparities in reading performance within schools. Lasso regression results suggest that almost half of teacher and school dimensions matter for within-school differences in reading performance between girls and boys when the overall population of 15-year-old students, teachers and principals surveyed within the TALIS-PISA link are considered (Figure 4.7). The eight dimensions highlighted by lasso regression analysis encompass teacher and school dimensions as well as factors with direct and indirect effects on student achievement. These dimensions are: teachers' classroom practices, teachers' motivation to join the profession, teachers' self-efficacy, school culture, teachers' participation in professional development activities of certain type and content, teacher collaboration, teachers' employment status and school leadership.

Figure 4.6. Gender disparities in student achievement within schools

Percentage of schools, by type of within-school gender disparity and by subject



1. Single-gender schools (i.e. all students surveyed in the school are same-gender students).
Notes: Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys. The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Viet Nam.
Countries and economies are ranked in descending order of the percentage of schools characterised by a statistically significant positive difference in performance between girls and boys.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.31, 4.32 and 4.33.

StatLink https://doi.org/10.1787/888934224125

Figure 4.7. Teacher and school factors that matter for within-school disparities in reading performance between girls and boys, based on lasso

Dimensions and variables selected by lasso regression based on responses of the overall population of 15-year-old students, teachers and principals surveyed within the TALIS-PISA link

Dimensions	Variables selected	Within-school differences in PISA score in reading between girls and boys[1] (sign of standardised coefficients: +/-)
Teacher dimensions with a direct effect on student achievement		
Classroom practices	Teachers' perceived disciplinary problems *(higher values indicate higher need for classroom discipline)*	+
Teacher dimensions with an indirect effect on student achievement		
Teacher characteristics	*No variables selected*	
Motivation to join the profession	Personal utility value	+
	Teaching as a first career choice	+
Initial education and training	*No variables selected*	
Well-being and job satisfaction	*No variables selected*	
Self-efficacy	Self-efficacy in instruction	-
Working hours	*No variables selected*	
School dimensions with a direct effect on student achievement		
Classroom characteristics (*classmates' characteristics and class size*)	*No variables selected*	
School culture	Collaborative school culture	+
	Stakeholder (i.e., parents and local community) involvement in school	-
School dimensions with an indirect effect on student achievement		
Induction	*No variables selected*	
Mentoring	*No variables selected*	
Professional development	Type: Reading professional literature	-
	Content: Teaching in a multicultural or multilingual setting	+
Feedback	*No variables selected*	
Collaboration	Professional collaboration	+
School innovativeness	*No variables selected*	
Employment status	Full-time (more than 90% of full time hours)	-
Formal appraisal	*No variables selected*	
School leadership	Principals' leadership activities: Observed instruction in the classroom	+

1. Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys.
Notes: Country fixed effects and student characteristics averaged at the school level, such as share of female students, share of students with an immigrant background and index of economic, social and cultural status, are always included among the variables selected by lasso.
Teacher variables are averaged for all teachers within the school.
Dimensions that are not selected are highlighted in light grey. Since lasso is applied as a model selection technique, it does not provide the standard errors required for performing statistical inference. The interpretation of the estimated standardised coefficients is conditional on the selected model and cannot be interpreted as causal. Moreover, in the presence of correlated explanatory variables, the signs of the coefficient estimates can swing based on which other independent variables are in the model.
For additional information on the full list of potential variables included in the lasso regressions, as well as more information on lasso in general, see Annex B.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 4.35.

Besides lasso regressions, one can also select the teacher and school dimensions that matter (the most) for within-school disparities in reading performance between girls and boys by identifying the dimensions that explain the highest shares of variance in within-school disparities. Although this approach cannot highlight the specific factors within a dimension that may matter the most for student achievement, it complements the findings of the lasso regression by revealing the relative importance of each teacher and school dimension in explaining the differences in within-school gender disparities in student achievement.

However, it is important to note that the shares of variance explained by each dimension may be artificially driven by the number of variables included in a given dimension. The dimensions that have the lowest explanatory power, such as school innovativeness, collaboration, mentoring, employment status and teacher characteristics, include four or fewer variables, while the dimensions that explain larger shares of the differences in school average performances include between seven and 14 variables (see more information in Annex B). Moreover, the observed explanatory power of the dimensions may also be inflated due to the low sample sizes. Thus, caution is warranted when interpreting these results.

Yet, with the aforementioned advantages and limitations in mind, it is possible to identify teacher and school dimensions that explain a substantial part of the differences in gender disparities in student performance within schools. The teacher and school dimensions that explain more than 20% of the variation in within-school gender gaps include teachers' participation in professional development activities of certain type (25%) and content (35%), teachers' initial teacher education, both in terms of content (35%) and sense of preparedness (22%), classroom characteristics (29%), teachers' classroom practices (29%), teachers' use of working time (29%), school leadership (26%) and teachers' formal appraisal (24%) (Figure 4.8).

In total, 12 dimensions are highlighted by either lasso or their explanatory power or both. Yet, based on the country-level regression results by each dimension taken separately, only five of these dimensions show a clear cross-country pattern in relation to gender disparities in reading performance within schools.[11] These include teachers' classroom practices, teachers' motivation to join the profession, school culture, teachers' formal appraisal and school leadership. The next section examines these five teacher and school dimensions more in detail at the country level. It is important to stress that, similarly to the lasso regression results, causal interpretation of the country-level regressions is not possible. All the results presented below are correlational and should be interpreted accordingly.

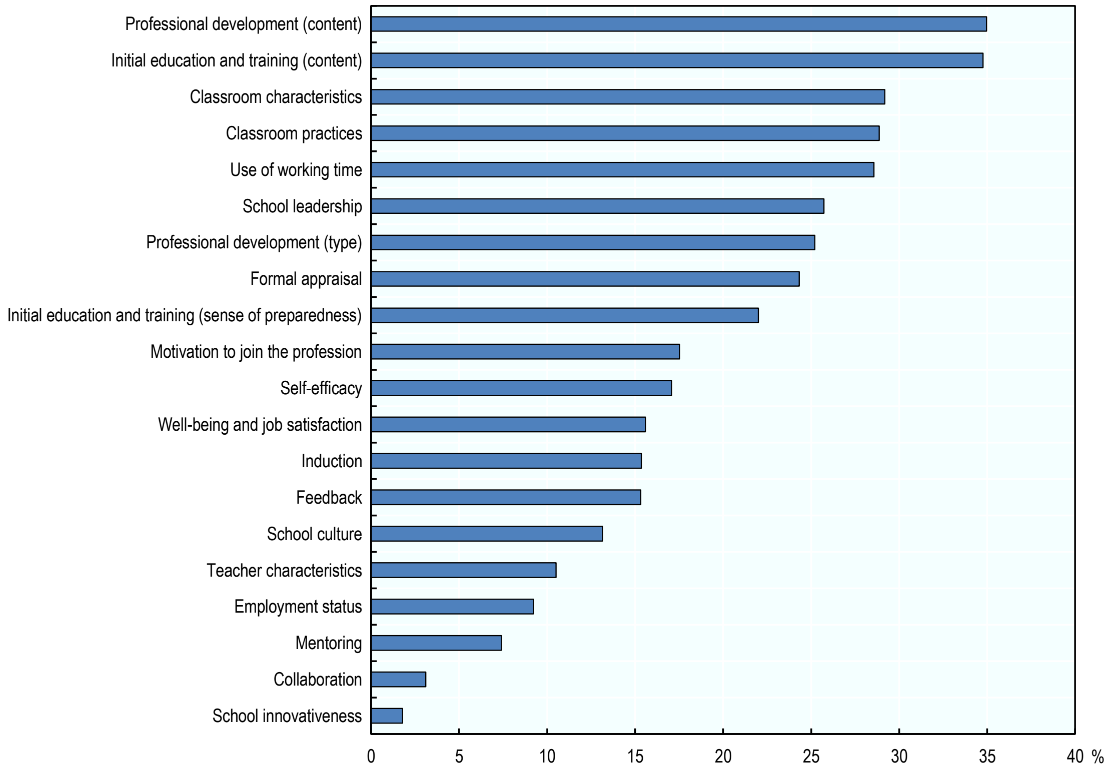
Teachers' classroom practices

Teachers' classroom practices are an important factor to consider in relation to gender gaps in student reading performance. They are highlighted by the lasso regression and explain a significant share (29%) of variance in within-school gender disparities in reading performance (Figure 4.7 and Figure 4.8). Country-level regression results reveal that two aspects of teachers' classroom practices stand out in their relation to gender gaps in reading achievement: teachers' perception of classroom disciplinary climate and the frequency with which teachers administer their own assessment to students.

First, as the extent of disciplinary issues perceived by the average school teacher increases, the greater the difference in PISA reading scores between girls and boys within schools is – favouring girls – on average across the countries and economies participating in the TALIS-PISA link, as well as in the Czech Republic, Georgia and Turkey (Figure 4.9). Boys might be more affected by a deteriorated classroom climate as they tend to be exposed to greater peer pressure than girls. They may also be more likely to disturb classes. Therefore, disciplinary issues may be more detrimental to boys' performance compared to girls'. These results also suggest that teachers' classroom management skills may matter for mitigating the gap in reading performance between girls and boys within schools.

Figure 4.8. Differences in within-school gender disparities in student achievement explained by teacher and school factors

Percentage of variance in within-school differences in the PISA reading score between girls and boys; explained by each dimension (TALIS-PISA link average)

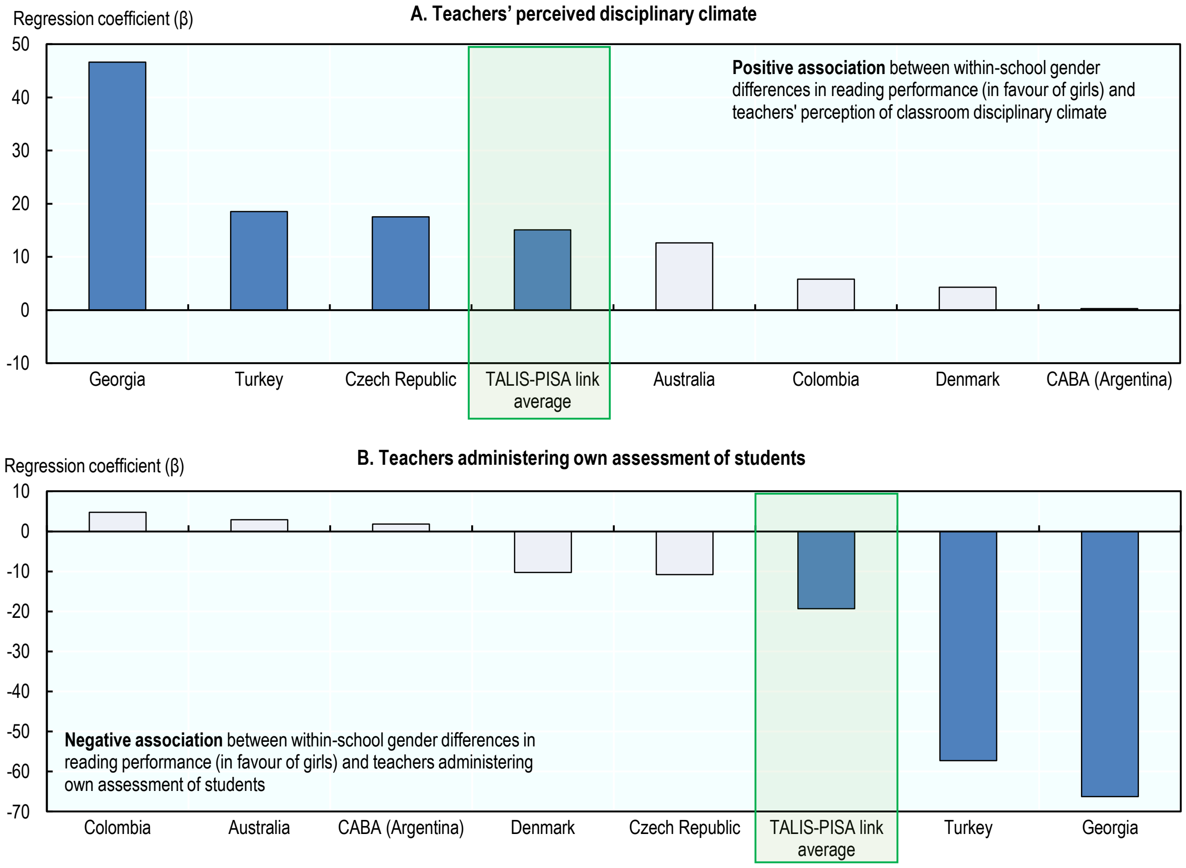


Notes: Teacher variables are averaged for all teachers within the school. Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys.
The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Malta and Viet Nam.
Values are ranked in descending order of the percentage of variance in within-school differences in the PISA reading score between girls and boys, explained by each dimension.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 4.36.

StatLink https://doi.org/10.1787/888934224144

Figure 4.9. Relationship between within-school gender disparities in reading and teachers' classroom practices

Change in within-school gender disparities in the PISA reading score associated with teachers' classroom practices



Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for the following classroom practices of teachers: class time spent on actual teaching and learning, teachers' autonomy over planning and teaching, use of practices related to clarity of instruction, use of practices related to cognitive activation, use of other assessment practices, such as providing written feedback on student work in addition to marking, letting students evaluate their own progress and observing students when working on particular tasks and providing immediate feedback; and for the following student characteristics averaged at the school level: gender, immigrant background and index of economic, social and cultural status.
Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys. Teacher variables are averaged for all teachers within the school.
The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Malta and Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in within-school gender disparities in the PISA reading score associated with teachers' classroom practices.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 4.37.

StatLink https://doi.org/10.1787/888934224163

Second, the more often teachers evaluate their students with their own assessment, the smaller the difference in reading performance reading between girls and boys, with boys enjoying the advantage. This holds true on average across all countries as well as in Georgia and Turkey (Figure 4.9). This finding suggests that boys benefit from more regular testing as it allows them to better self-regulate and focus on schoolwork. Research shows that a learning-oriented environment is especially beneficial for boys (Legewie and DiPrete, 2012[53]). Such an environment can shape how masculinity is constructed in peer culture among boys and, thereby, influence boys' orientation toward school. Yet, it is also possible that girls benefit less from more frequent assessment by teachers due to their increased anxiety about frequent testing and induced competition (Devine et al., 2012[54]; Gneezy, Niederle and Rustichini, 2003[55]; McLean and Anderson, 2009[56]; Niederle and Vesterlund, 2011[57]), even though they may be better at sustaining their performance during longer tests (Balart and Oosterveen, 2019[58]). PISA 2018 results show that, in the majority of participating countries and economies, boys were more likely to express more positive attitudes towards competition than girls (OECD, 2019[7]).

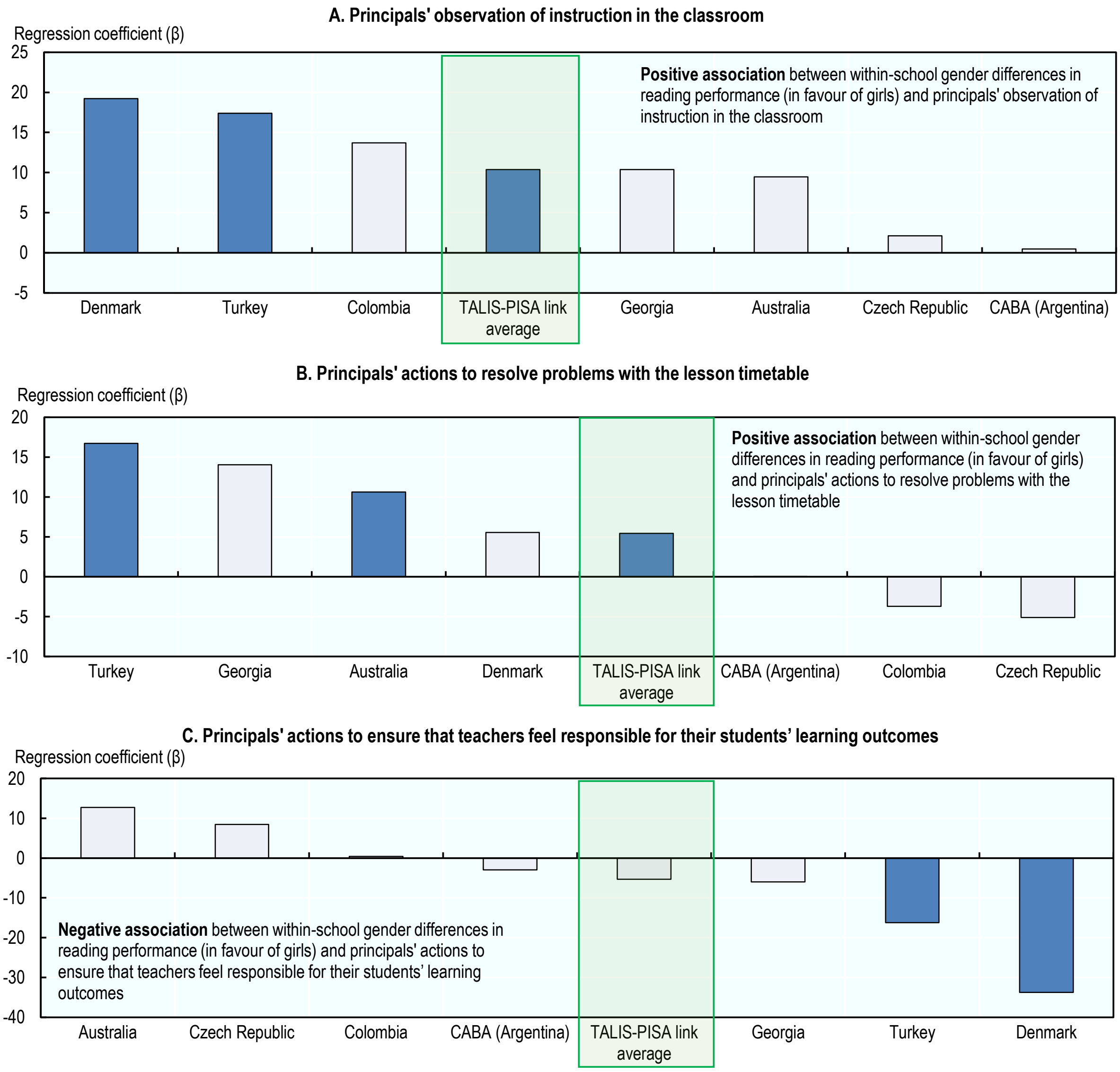
School leadership

Findings of lasso regression analysis presented above highlighted school leadership, in particular the frequency with which principals reported observing instruction in the classroom as an important predictor for within-school differences in reading performance between girls and boys (Figure 4.7). School leadership was also highlighted above as a dimension that explains a substantial share (26%) of the variation in gender-related disparities in student achievement (Figure 4.8). Country-level regression results are in line with these findings. On average across participating countries and economies as well as in Denmark and Turkey, schools whose principals tend to observe instruction in the classroom more often are characterised by larger differences in reading performance in favour of girls (Figure 4.10). This finding suggests that, compared to boys, girls may gain more from direct forms of instructional leadership activities such as observing instruction in the classroom. This may be the case, for instance, if girls find extra motivation to focus on schoolwork in schools where there is extended control and oversight from the school leader on what happens in classrooms. Nevertheless, this finding might also point to a reverse causal relation, with school leaders engaging more in instructional leadership activities in classes and schools with a higher share of low-achieving boys.

Country-level regression results also suggest that two other activities of school leaders present some association with gender gap in achievement: dealing with the school timetable and fostering teachers' feeling of responsibility for their students' achievement. First, the more often school leaders engage in resolving problems with the lesson timetable in the school, the larger the outperformance of girls over boys in reading performance. This holds true on average across countries and economies as well as in Australia and Turkey (Figure 4.10). This finding suggests that boys are more disturbed than girls in their learning by practical organisational issues occurring in the school. This corroborates, once again, the fact that girls are more able to stay focused than boys,[12] regardless of problems arising. This echoes several research studies suggesting that, for many boys, it is not acceptable to be seen as interested in school work (DiPrete and Buchmann, 2013[39]). Boys adopt attitudes that include a disregard for authority, academic work and formal achievement (Salisbury, Rees and Gorard, 1999[59]), which might express more acutely when there are practical and organisational disruptions in the school. Second, the more school leaders ensure that teachers feel responsible for their students' learning outcomes, the better boys perform in reading compared to girls. This holds true in Denmark and Turkey (Figure 4.10). This kind of instructional leadership might be particularly efficient in motivating teachers to support boys and, in particular, low-achieving boys. It might also be the case that boys are more stimulated than girls in a school environment when a significant weight is given to performance and accountability. In fact, this resonates with the finding presented above about boys benefiting more from more regular testing – a school environment where accountability manifests in frequent testing may be especially beneficial for boys. It may allow them to better self-regulate and focus on schoolwork as boys also tend to express more positive attitudes towards competition than girls.

Figure 4.10. Relationship between within-school gender disparities in reading and school leadership

Change in within-school gender disparities in the PISA reading score associated with principals' leadership activities



Notes: Results of linear regression based on responses of principals and 15-year-old students. Controlling for the following leadership activities of principals: collaborating with teachers to solve classroom discipline problems, providing feedback to teachers, supporting co-operation among teachers to develop new teaching practices, ensuring that teachers take responsibility for improving their teaching skills, providing parents or guardians with information on the school and student performance, reviewing school administrative procedures and reports, collaborating with principals from other schools on challenging work tasks and working on a professional development plan for the school; and for the following student characteristics averaged at the school level: gender, immigrant background and index of economic, social and cultural status.
Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys. Teacher variables are averaged for all teachers within the school.
The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Malta and Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in within-school gender disparities in the PISA reading score associated with principals' leadership activities.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 4.38.

StatLink https://doi.org/10.1787/888934224182

School culture

School culture explains a relatively low share of the variation in within-school gender differences, i.e. only 13% on average across participating countries and economies (Figure 4.8). Yet, school culture is selected by lasso as a key predictor of gender gaps in reading. In particular, having a collaborative school culture and stakeholders such as parents and local community involved in school-related activities are significantly associated with gender disparities in reading performance within schools (Figure 4.7). Country-level regression results show consistent cross-country patterns for certain indicators of school culture. This is the case for stakeholder involvement in school as well as teacher-student relations, which both seem to matter for school-level gender differences in reading performance (Table 4.39). Notably, in Colombia and Denmark, within-school disparities in reading performance between girls and boys tend to decrease in favour of boys in those schools where stakeholders (i.e. parents and community) are involved in school-related activities (Figure 4.11). Similarly, in schools where the average school teacher consider teacher-student relationships to be positive, gender disparities in reading performance are smaller, in favour of boys, in Australia, Colombia and on average across participating countries and economies (Figure 4.11). These findings suggest that girls may be better at self-regulating and staying focused on their school work while boys may need more support from their parents. They may also benefit more from a positive relationship with their teachers.

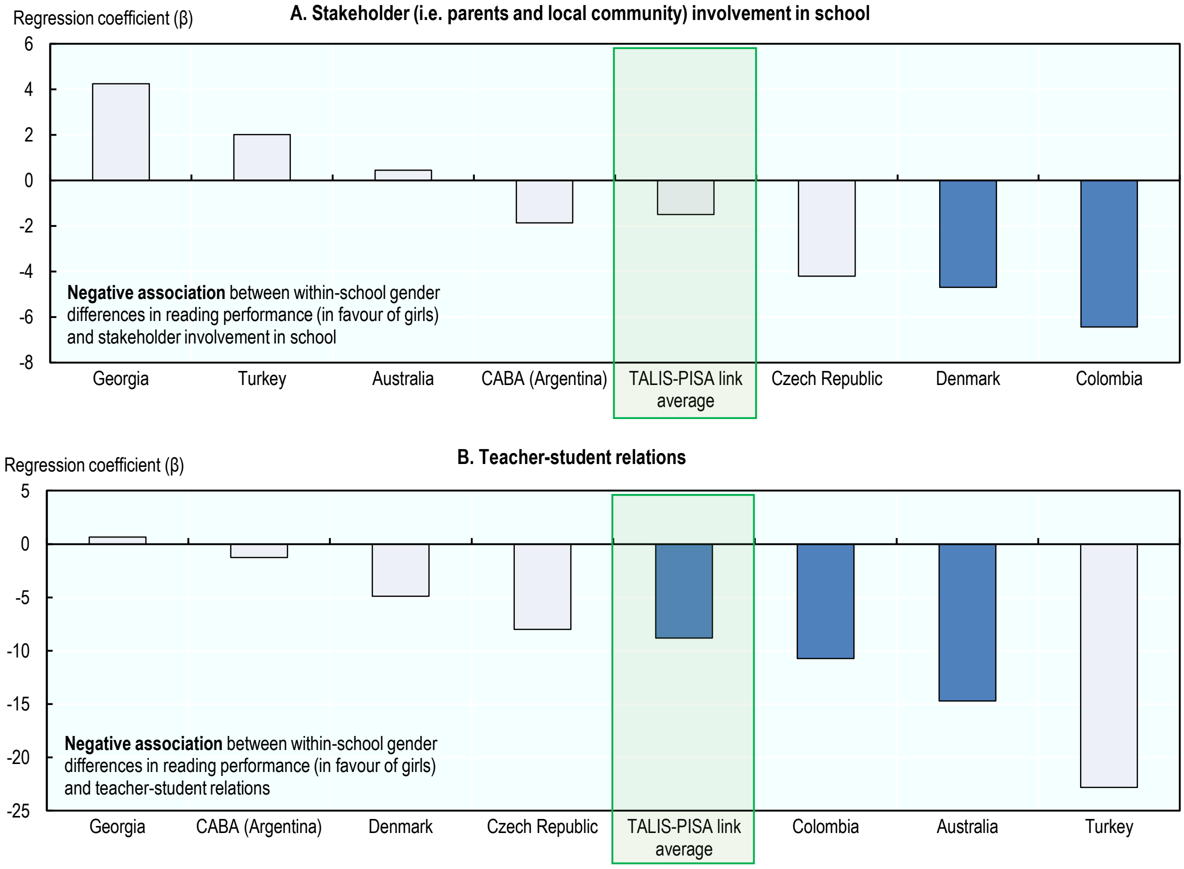
Teachers' formal appraisal

The dimension of formally appraising teachers is not selected by lasso as an important predictor for within-school gender disparities in reading performance (Figure 4.7). However, it explains a substantial part (24%) of the variation in school-level gaps in PISA reading scores between girls and boys (Figure 4.8). Country-level analysis reveals a significant association between features of teachers' formal appraisal and school-level gender differences in reading performance that holds in various countries. Notably, the more external individuals and bodies formally appraise teachers in the school, the more likely boys will perform as well as (if not better than) girls in reading on average across participating countries and economies as well as in Denmark, Georgia and Turkey (Table 4.40). This finding indicates that formal evaluation of teachers may not only be an important element of high-performing schools, especially when it has a strong focus on teachers' continuous improvement (OECD, 2018[60]), but it can also address gender disparities in student achievement. Teacher appraisal can encourage teachers to reflect on and improve their teaching practices. Improved classroom management practices, for example, may benefit boys more than girls.

External teacher appraisal sources can be perceived as more objective and less judgemental than appraisal by the school management team or other colleagues who work in the same school (OECD, 2020[26]). It can also happen in a wide variety of forms involving very different individuals or bodies. Thus, it is difficult to draw general conclusions regarding formal evaluation of teachers by external sources. The association found between teachers' formal appraisal and within-school gender disparities in reading performance may simply point to the importance of formal appraisal in mitigating gender disparities in student achievement, irrespective of the source of the teacher evaluation.

Figure 4.11. Relationship between within-school gender disparities in reading and school culture

Change in within-school gender disparities in the PISA reading score associated with school culture



Notes: Results of linear regression based on responses of principals, teachers and 15-year-old students. Controlling for the following aspects of school culture: collaborative school culture and teachers' actions towards achieving academic excellence; and for the following student characteristics averaged at the school level: gender, immigrant background and index of economic, social and cultural status.
Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys.
Teacher variables are averaged for all teachers within the school.
The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Malta and Viet Nam.
Statistically significant coefficients are marked in a darker tone (see Annex B).
Countries and economies are ranked in descending order of the change in within-school gender disparities in the PISA reading score associated with school culture.
Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Table 4.39.

StatLink https://doi.org/10.1787/888934224201

Teachers' motivation to join the profession

Teachers' motivation to join the profession explains a relatively low share (18%) of the variation in within-school gender differences on average across participating countries and economies (Figure 4.8). Nevertheless, this teacher factor is selected by lasso. In particular, two of its indicators – the share of teachers for whom teaching was a first career choice as well as the weight given by the school's teachers to personal-utility motivations when they considered joining the profession – have a significant relationship with school-level differences in reading performance between girls and boys (Figure 4.7). Yet, the country-level regression results do not confirm these findings from lasso. Nevertheless, the relation between gender gaps within schools and the share of teachers who chose teaching as their first career holds both for the overall population surveyed within the TALIS-PISA link (through lasso) and across countries. As the share of teachers for whom teaching was a first career choice increases within schools, the differences in PISA reading scores between girls and boys tend to increase in favour of girls on average across countries as well as in the Czech Republic, Denmark and Turkey (Table 4.41). As shown in TALIS 2018, teachers whose first career choice was teaching are more likely to be satisfied with their job in almost all countries and economies participating in TALIS, and they also tend to report higher self-efficacy (OECD, 2019[28]). The combination of all these factors might be driving larger gender disparities in reading in favour of girls. Yet, the reasons that explain this finding are unknown, warranting the need for further research.

Teacher gender and attitudes in relation to gender disparities in student achievement

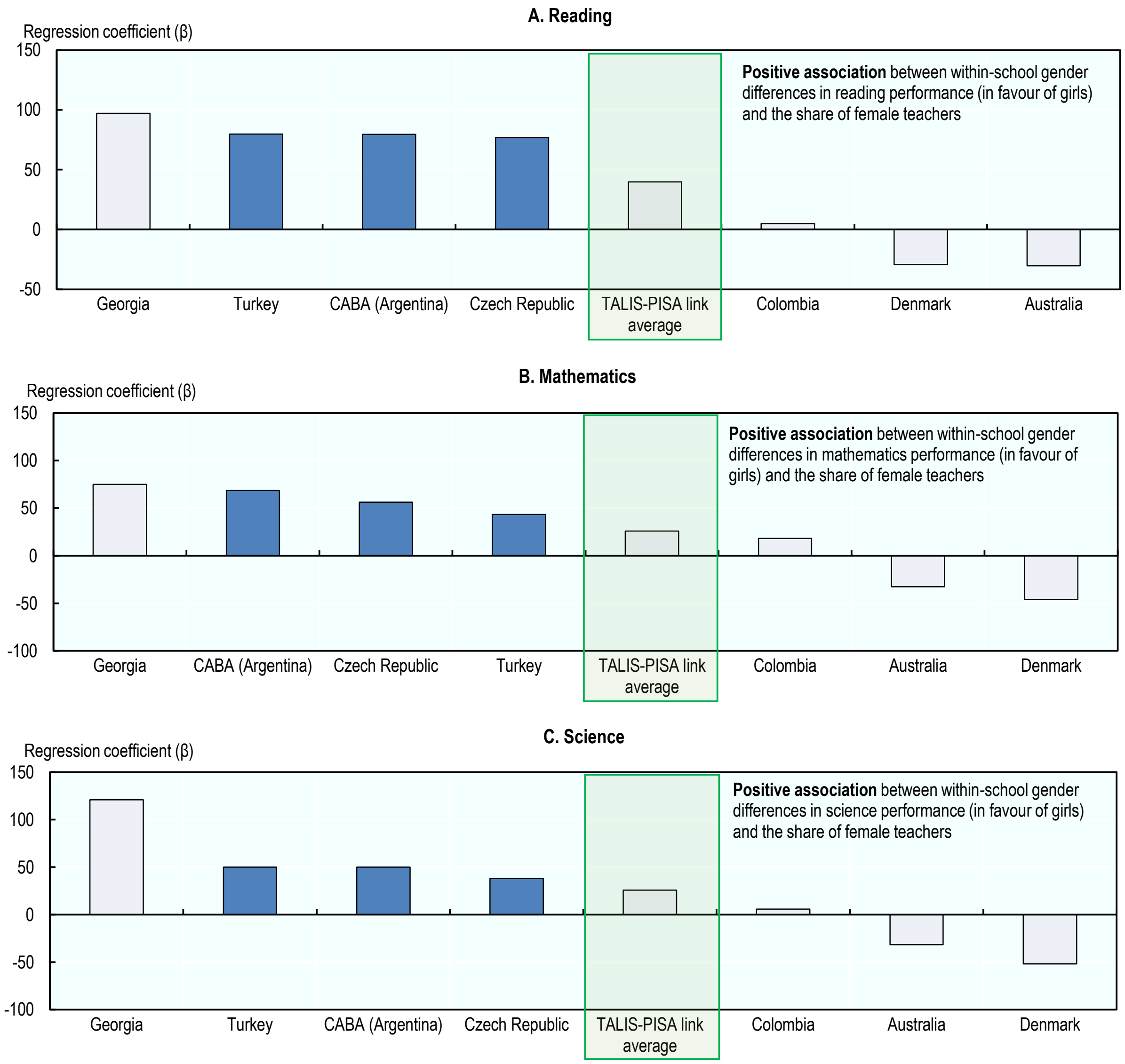
In the quest for teacher and school factors that can play a role in mitigating gender disparities in student performance, this section aims to test specific hypotheses about system-level policies and teachers' attitudes in relation to gender disparities in student performance in all three subjects (reading, mathematics and science). Do female teachers act as role models for girls, especially when it comes to closing the gender gap in mathematics or science? In the same way, can male teachers act as role models for boys to close the gender gap in reading performance? Do teachers' attitudes matter for gender disparities in student achievement? The TALIS principal questionnaire includes a question that focuses on teachers' attitudes towards gender discrimination, making it possible to examine whether teachers' beliefs can help to address gender gaps in student achievement.

TALIS-PISA link data suggest that teachers may act as role models for students of their gender. In CABA (Argentina), the Czech Republic and Turkey, as the share of female teachers increases in the schools, the differences in PISA scores between girls and boys also tend to increase in reading performance and decrease for mathematics and science in favour of girls (Figure 4.12). In the same way, as the share of male teachers increases in the school, the better boys perform compared to girls in all subjects. These results hold while accounting for the average school teacher's years of teaching experience. These findings suggest that, at least in certain education systems, teachers may have a bigger influence on students of their gender either because of distinct teaching practices that suit same-gender students better or by overcoming gender stereotypes because they are role models for these students.

The TALIS principal questionnaire has a specific question that asks school leaders about their views on teachers' attitudes regarding gender discrimination. Notably, it asks principals whether "none or almost none", "some", "many" or "all or almost all" teachers in the school would agree that students should learn how to avoid gender discrimination; it is important to treat female and male students equally. Country-level regressions results are not significant for most participating countries and economies (Tables 4.45, 4.46 and 4.47). Yet, in schools in Turkey where principals perceive that all or almost all teachers in the school would agree that it is important to treat female and male students equally, within-school gender disparities in student performance tend to increase in favour of girls. This result holds for reading, mathematics and science and it aligns with past research showing that the prevalence of gender stereotypes and gender gaps are closely related (Else-Quest, Hyde and Linn, 2010[42]; Guiso et al., 2008[43]; Nollenberger, Rodríguez-Planas and Sevilla, 2016[44]; Nosek, Banaji and Greenwald, 2002[45]).

Figure 4.12. Relationship between within-school gender disparities in student achievement and teachers' gender

Change in within-school gender disparities in PISA scores associated with teachers' gender, by subject



Notes: Results of linear regression based on responses of 15-year-old students and teachers. Controlling for teachers' average years of teaching experience; and for the following student characteristics averaged at the school level: gender, immigrant background and index of economic, social and cultural status.

Within-school differences in performance between girls and boys are defined as the school-level average PISA score for girls minus the school-level average PISA score for boys. Differences are positive when they are in favour of girls and negative when they are in favour of boys.

Teacher variables are averaged for all teachers within the school.

The TALIS-PISA link average corresponds to the arithmetic mean of the estimates of participating countries and economies, excluding Malta and Viet Nam.

Statistically significant coefficients are marked in a darker tone (see Annex B).

Countries and economies are ranked in descending order of the change in within-school gender disparities in PISA scores associated with teachers' gender.

Sources: OECD, TALIS 2018 Database; OECD, PISA 2018 Database, Tables 4.42, 4.43 and 4.44.

StatLink https://doi.org/10.1787/888934224220

Summary

Two main sets of conclusions can be drawn from the findings based on country-level analyses of gender gaps in reading performance.

First, boys seem to be more disturbed than girls by classroom disciplinary problems and school organisational issues. Indeed, as the level of disciplinary issues perceived by the average school teacher increases, the difference in PISA reading scores between girls and boys within schools tends to increase further in favour of girls. In addition, the more often school leaders observe instruction in classrooms (most likely as a remedial measure) and resolve problems with the lesson timetable in the school, the larger the outperformance of girls over boys in reading. Boys might be more affected by deteriorated learning conditions as they tend to be exposed to greater peer pressure than girls. Boys may also be more likely to be the students disturbing lessons. Overall, they seem less able than girls to stay focused on their schoolwork when disciplinary or practical issues arise in school. The findings suggest potential measures to help boys close the gap with girls in reading: greater involvement and support from parents in school-related activities as well as positive relationships between teachers and students. This suggests that 15-year old boys, more than girls, need support from significant adults such as their parents and teachers to self-regulate and be achievement-focused.

Second, boys are more likely to perform as well as (or even better) than girls in reading in schools where a culture of student assessment, teacher accountability and appraisal prevails. The more often teachers evaluate their students by administering their own assessment, the smaller the difference in reading performance between girls and boys to the advantage of boys. This may suggest that boys benefit from more regular testing as it allows them to better self-regulate and focus on schoolwork. Yet, this might also be due to girls' increased anxiety about frequent testing and induced competition. In addition, the more school leaders ensure that teachers feel responsible for their students' learning outcomes and the more often teachers are formally appraised by external individuals and bodies, the better boys perform in reading compared to girls. This suggests that teacher accountability and appraisal could help to address gender disparities in student achievement. This could be the case, in particular, if teachers are given opportunities to reflect on their teaching practice and find ways to better support low- and middle-achievers among which boys are over-represented in reading.

References

Ainley, J. and R. Carstens (2018), "Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework", *OECD Education Working Papers*, No. 187, OECD Publishing, Paris, https://dx.doi.org/10.1787/799337c2-en. [25]

Autor, D. et al. (2016), "School quality and the gender gap in educational achievement", *American Economic Review*, Vol. 106/5, pp. 289-295, http://dx.doi.org/10.1257/aer.p20161074. [37]

Balart, P. and M. Oosterveen (2019), "Females show more sustained performance during test-taking than males", *Nature Communications*, Vol. 10/3798, https://doi.org/10.1038/s41467-019-11691-y. [58]

Breda, T. et al. (2020), "Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools", *PSE Working Papers*, No. 2018-06, HAL Sciences de l'Homme et de la Société, https://halshs.archives-ouvertes.fr/halshs-01713068v3. [48]

Brenøe, A. and S. Lundberg (2018), “Gender gaps in the effects of childhood family environment: Do they persist into adulthood?”, *European Economic Review*, Vol. 109, pp. 42-62, http://dx.doi.org/10.1016/j.euroecorev.2017.04.004. [38]

Burke, M. and T. Sass (2013), “Classroom peer effects and student achievement”, *Journal of Labor Economics*, Vol. 31/1, pp. 51-82, http://dx.doi.org/10.1086/666653. [31]

Cameron, A. and P. Trivedi (2009), *Microeconometrics Using Stata*, Stata Press, College Station, TX. [14]

Carlana, M. (2019), “Implicit stereotypes: Evidence from teachers’ gender bias”, *The Quarterly Journal of Economics*, Vol. 134/3, pp. 1163-1224, http://dx.doi.org/10.1093/qje/qjz008. [51]

Carrell, S., M. Page and J. West (2010), “Sex and science: How professor gender perpetuates the gender gap”, *The Quarterly Journal of Economics*, Vol. 125/3, pp. 1101-1144, http://dx.doi.org/10.1162/qjec.2010.125.3.1101. [47]

Cimpian, J. et al. (2016), “Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts”, *AERA Open*, Vol. 2/4, pp. 1-19, http://dx.doi.org/10.1177/2332858416673617. [50]

Dee, T. (2005), “A teacher like me: Does race, ethnicity, or gender matter?”, *American Economic Review*, Vol. 95/2, pp. 158-165, http://dx.doi.org/10.1257/000282805774670446. [46]

Devine, A. et al. (2012), “Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety”, *Behavioral and Brain Functions*, Vol. 8/33, pp. 1-9, http://dx.doi.org/10.1186/1744-9081-8-33. [54]

DiPrete, T. and C. Buchmann (2013), *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*, Russell Sage Foundation, New York, NY. [39]

Else-Quest, N., J. Hyde and M. Linn (2010), “Cross-national patterns of gender differences in mathematics: A meta-analysis.”, *Psychological Bulletin*, Vol. 136/1, pp. 103-127, http://dx.doi.org/10.1037/a0018053. [42]

Erikson, R. et al. (2005), “On class differentials in educational attainment”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102/27, pp. 9730-9733, http://dx.doi.org/10.1073/pnas.0502433102. [3]

Gneezy, U., M. Niederle and A. Rustichini (2003), “Performance in competitive environments: Gender differences”, *The Quarterly Journal of Economics*, Vol. 118/3, pp. 1049-1074, http://dx.doi.org/10.1162/00335530360698496. [55]

Gromada, A. and C. Shewbridge (2016), “Student Learning Time: A Literature Review”, *OECD Education Working Papers*, No. 127, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jm409kqqkjh-en. [24]

Guiso, L. et al. (2008), “Culture, gender, and math”, *Science*, Vol. 320/5880, pp. 1164-1165, http://dx.doi.org/10.1126/science.1154094. [43]

Hadjar, A. et al. (2014), “Gender and educational achievement”, *Educational Research*, Vol. 56/2, pp. 117-125, http://dx.doi.org/10.1080/00131881.2014.898908. [40]

Hanushek, E. et al. (2015), "Returns to skills around the world: Evidence from PIAAC", *European Economic Review*, Vol. 73, pp. 103-130, http://dx.doi.org/10.1016/j.euroecorev.2014.10.006. [1]

Hattie, J. (2009), *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*, Routledge, London. [20]

Koenker, R. (2017), "Quantile regression: 40 years on", *Annual Review of Economics*, Vol. 9/1, pp. 155-176, http://dx.doi.org/10.1146/annurev-economics-063016-103651. [11]

Koenker, R. (2005), *Quantile Regression*, Cambridge University Press, Cambridge, http://dx.doi.org/10.1017/CBO9780511754098. [12]

Koenker, R. and G. Bassett (1978), "Regression quantiles", *Econometrica: The Journal of the Econometric Society*, Vol. 46/1, pp. 33-50, http://dx.doi.org/10.2307/1913643. [13]

Lavy, V. and E. Sand (2018), "On the origins of gender gaps in human capital: Short- and long-term consequences of teachers' biases", *Journal of Public Economics*, Vol. 167, pp. 263-279, http://dx.doi.org/10.1016/j.jpubeco.2018.09.007. [49]

Lavy, V., O. Silva and F. Weinhardt (2012), "The good, the bad, and the average: Evidence on ability peer effects in schools", *Journal of Labor Economics*, Vol. 30/2, pp. 367-414, http://dx.doi.org/10.1086/663592. [32]

Le Donné, N., P. Fraser and G. Bousquet (2016), "Teaching Strategies for Instructional Quality: Insights from the TALIS-PISA Link Data", *OECD Education Working Papers*, No. 148, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jln1hlsr0lr-en. [21]

Legewie, J. and T. DiPrete (2012), "School context and the gender gap in educational achievement", *American Sociological Review*, Vol. 77/3, pp. 463-485, http://dx.doi.org/10.1177/0003122412440802. [53]

Lounkaew, K. (2013), "Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy data", *Economics of Education Review*, Vol. 37, pp. 213-225, http://dx.doi.org/10.1016/j.econedurev.2013.09.003. [19]

McLean, C. and E. Anderson (2009), "Brave men and timid women? A review of the gender differences in fear and anxiety", *Clinical Psychology Review*, Vol. 29/6, pp. 496-505, http://dx.doi.org/10.1016/j.cpr.2009.05.003. [56]

Mostafa, T., A. Echazarra and H. Guillou (2018), "The science of teaching science: An exploration of science teaching practices in PISA 2015", *OECD Education Working Papers*, No. 188, OECD Publishing, Paris, https://dx.doi.org/10.1787/f5bd9e57-en. [17]

Muijs, D. et al. (2014), "State of the art: Teacher effectiveness and professional learning", *School Effectiveness and School Improvement*, Vol. 25/2, pp. 231-256, http://dx.doi.org/10.1080/09243453.2014.885451. [22]

Muijs, D. and D. Reynolds (2001), *Effective Teaching: Evidence and Practice*, Sage Publications, London. [29]

Niederle, M. and L. Vesterlund (2011), "Gender and competition", *Annual Review of Economics*, Vol. 3, pp. 601-630, http://dx.doi.org/10.1146/annurev-economics-111809-125122. [57]

Nollenberger, N., N. Rodríguez-Planas and A. Sevilla (2016), "The math gender gap: The role of culture", *American Economic Review*, Vol. 106/5, pp. 257-261, http://dx.doi.org/10.1257/aer.p20161121. [44]

Nosek, B., M. Banaji and A. Greenwald (2002), "Math = male, me = female, therefore math ≠ me", *Journal of Personality and Social Psychology*, Vol. 83/1, pp. 44-59, http://dx.doi.org/10.1037/0022-3514.83.1.44. [45]

OECD (2020), *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/19cf08df-en. [26]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b5fd1b8f-en. [7]

OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/1d0bc92a-en. [28]

OECD (2018), *Effective Teacher Policies: Insights from PISA*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264301603-en. [60]

OECD (2018), *Equity in Education: Breaking Down Barriers to Social Mobility*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264073234-en. [5]

OECD (2016), *Low-Performing Students: Why They Fall Behind and How To Help Them Succeed*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264250246-en. [10]

OECD (2016), *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264258051-en. [2]

OECD (2015), *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264229945-en. [8]

OECD (2014), *TALIS 2013 Results: An International Perspective on Teaching and Learning*, TALIS, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264196261-en. [27]

OECD (2010), *The High Cost of Low Educational Performance: The Long-run Economic Impact of Improving PISA Outcomes*, PISA, OECD, http://dx.doi.org/10.1787/9789264077485-en. [9]

Parker, P., B. Van Zanden and R. Parker (2018), "Girls get smart, boys get smug: Historical changes in gender differences in math, literacy, and academic social comparison and achievement", *Learning and Instruction*, Vol. 54, pp. 125-137, http://dx.doi.org/10.1016/j.learninstruc.2017.09.002. [52]

Patall, E., H. Cooper and A. Allen (2010), "Extending the school day or school year: A systematic review of research (1985-2009)", *Review of Educational Research*, Vol. 80/3, pp. 401-436, http://dx.doi.org/10.3102/0034654310377086. [23]

Raitano, M. and F. Vona (2013), "Peer heterogeneity, school tracking and students' performances: Evidence from PISA 2006", *Applied Economics*, Vol. 45/32, pp. 4516-4532, http://dx.doi.org/10.1080/00036846.2013.791020. [16]

Rangvid, B. (2007), "School composition effects in Denmark: Quantile regression evidence from PISA 2000", *Empirical Economics*, Vol. 33, pp. 359-388, http://dx.doi.org/10.1007/s00181-007-0133-6. [15]

Robinson-Cimpian, J. et al. (2014), “Teachers’ perceptions of students’ mathematics proficiency may exacerbate early gender gaps in achievement”, *Developmental Psychology*, Vol. 50/4, pp. 1262-1281, http://dx.doi.org/10.1037/a0035073. [41]

Roemer, J. (1998), *Equality of Opportunity*, Harvard University Press, Cambridge, MA. [6]

Rose, H. and J. Betts (2004), “The effect of high school courses on earnings”, *Review of Economics and Statistics*, Vol. 86/2, pp. 497-513, http://dx.doi.org/10.1162/003465304323031076. [4]

Sacerdote, B. (2011), “Peer effects in education: How might they work, how big are they and how much do we know thus far?”, in Hanushek, E., S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1. [30]

Salisbury, J., G. Rees and S. Gorard (1999), “Accounting for the differential attainment of boys and girls at school”, *School Leadership and Management*, Vol. 19/4, pp. 403-426, http://dx.doi.org/10.1080/13632439968943. [59]

Teng, Y. (2020), “The relationship between school climate and students’ mathematics achievement gaps in Shanghai China: Evidence from PISA 2012”, *Asia Pacific Journal of Education*, Vol. 40/3, pp. 356-372, http://dx.doi.org/10.1080/02188791.2019.1682516. [18]

UNESCO (2018), *Global Education Monitoring Report Gender Review 2018: Meeting our Commitments to Gender Equality in Education*, UNESCO, Paris, https://unesdoc.unesco.org/ark:/48223/pf0000261593. [36]

van Ewijk, R. and P. Sleegers (2011), “The Effect of Peer Socioeconomic Status on Student Achievement: A Meta-Analysis”, *SSRN Electronic Journal*, http://dx.doi.org/10.2139/ssrn.1402645. [33]

Wang, M. and J. Degol (2016), “School climate: A review of the construct, measurement, and impact on student outcomes”, *Educational Psychology Review*, Vol. 28/2, pp. 315-352, http://dx.doi.org/10.1007/s10648-015-9319-1. [34]

Wilder, S. (2014), “Effects of parental involvement on academic achievement: a meta-synthesis”, *Educational Review*, Vol. 66/3, pp. 377-397, http://dx.doi.org/10.1080/00131911.2013.780009. [35]

Notes

[1] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[2] Since Viet Nam does not have data on PISA test scores, it is not included in the analyses presented in Chapter 4.

[3] Analyses on socio-economic disparities in achievement within schools are available upon request to the authors of the report. Authors conducted analyses using two alternative definitions of socially disadvantaged and advantaged students. In the first approach, socially disadvantaged students were defined as those students who are in the bottom *quarter* of the PISA index of economic, social and cultural status (ESCS), and in the second approach those in the bottom *half*. Conversely, socially advantaged students are those in the top *quarter* of the PISA index of economic, social and cultural status, and in the second approach those in the top *half*. In the first approach, neither lasso regressions based on the overall population surveyed within the TALIS-PISA link, nor country-level regressions highlight teacher and school factors that may be significantly related to socio-economic disparities in student performance within schools. In the second approach, which is applied on an increased number of schools enrolling both types of students but implies, at the same, smaller socio-economic gaps within schools, lasso regressions do select quite a few dimensions as key predictors of socio-economic gaps in achievement within schools. However, the lasso findings do not align well with those from the country-level regressions, which are rather inconsistent. It was decided not to report these analyses, given their lack of robustness.

[4] Among the eight countries/economies with available performance data, four of them have too few schools that enrol both native and migrant students (Colombia, the Czech Republic, Georgia and Turkey). For the remaining four countries/economies, the number of schools with a significant within-school gap between migrant and non-migrant students is limited (especially in CABA [Argentina] and Malta).

[5] For instance, analysis based on TALIS data show that in the case of teachers' satisfaction with the profession, only a small percentage (i.e. 4%) of the total variance comes from differences between schools (OECD, 2020[26]).

[6] Based on findings from Rangvid (2007[15]), quantile regression results suggest differential school composition effects across the conditional reading score distribution, with students at the lower quantiles benefitting more from socially advantaged peers. Students at lower and higher quantiles of the conditional mathematics score distribution seem to benefit equally from attending schools with a better student intake. Finally, Rangvid's findings are at best marginally significant when it comes to the relationship between the school socio-economic composition and students' science test scores.

[7] The role of classmates' characteristics as confounding factors in the association between a teacher or school factor, and student achievement by level of performance was also explored. However, due to computational process (i.e. non-convergence of the quantile regression model), the estimates for the models including controls for classmates' characteristics were missing for certain countries. Therefore, the models that include controls for classmates' characteristics are neither reported nor commented on.

[8] Teacher gender bias was measured using class-gender differences in scores between school exams graded by teachers and national exams graded blindly by external examiners.

[9] Excluding Malta, where single-gender schools constitute around 60% of all schools, from the TALIS-PISA link average leads to a decrease in the share of same-gender schools without gender gap estimates of 6 percentage points from 18% to 12%.

[10] In Malta, there are only 17 out of the 44 schools surveyed that are not single-gender schools (i.e. all students surveyed in the school are same-gender students) and where the within-school differences in performance between girls and boys can be computed.

[11] Examining each of the remaining seven dimensions separately and more closely at the country-level – teachers' participation in professional development activities (both type and content), teachers' initial teacher education (both in terms of content and sense of preparedness), classroom characteristics, teachers' self-efficacy, teachers' employment status and teacher collaboration – shows very few significant results that would follow a consistent pattern across countries, hence provides little evidence for a significant relationship with school-level gender disparities in reading performance.

[12] For example, girls are overwhelmingly more likely than boys to spend time doing homework (OECD, 2015[8]).

Annex A. Technical notes on TALIS-PISA link data

Sampling procedures

The objective of the Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link, referred to as TALIS-PISA link[1] in the report, was to obtain, in each participating country and economy, representative samples of 15-year-old students, teachers teaching 15-year-old students and schools with students of this age that can be linked and analysed together. Thus, in each country and economy that opted to participate in the TALIS-PISA link option of TALIS 2018 (hereafter "TALIS-PISA link 2018"), a random sub-sample was drawn from the PISA sample of schools. The international sampling plan prepared for TALIS-PISA link 2018 and PISA 2018 used a stratified two-stage probability sampling design. This means that teachers and students (second stage units, or secondary sampling units) were to be randomly selected from the list of in-scope teachers and students in each of the randomly selected schools (first stage units, or primary sampling units). A more detailed description of the survey design and its implementation can be found in the TALIS 2018 Technical Report (OECD, 2019[1]) and the PISA 2018 Technical Report (OECD, 2020[2]).

The international target population of TALIS-PISA link 2018 restricts the survey to those teachers and principals who work in schools surveyed by PISA that provide instruction for 15-year-old students. Only teachers who teach regular classes to PISA-eligible students in ordinary schools surveyed by PISA are covered by the TALIS-PISA link. Teachers working in schools exclusively devoted to children with special needs are not part of the international target population and are deemed out of scope. Teachers working with special needs students in a regular school setting were considered in scope in TALIS-PISA link. However, when a school is made up exclusively of these teachers, the school itself is said to be out of scope. Teacher aides, pedagogical support staff (e.g. guidance counsellors and librarians) and health and social support staff (e.g. doctors, nurses, psychiatrists, psychologists, occupational therapists and social workers) were not considered to be teachers and, thus, not part of the TALIS-PISA link international target population.

In PISA 2018, the international target population of students includes those who are aged between 15 years and 3 months and 16 years and 2 months at the time of the 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. The PISA 2018 target population does not include residents of a country who attend school in another country. However, it does include foreign nationals who attend school in the country of assessment.

For national reasons, participating countries and economies could choose to reduce their coverage of the target population by excluding, for instance, a small, remote geographical region due to inaccessibility, or language differences, possibly due to political, organisational or operational reasons, or presence of special education needs students. However, efforts were made to keep these exclusions to a minimum (i.e. up to a total of 5% of the relevant teacher and student populations, either by excluding schools or excluding students and teachers within schools).

The following categories of schools were excluded from the sample:

- 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 exclusion of students", such as schools for the blind (i.e. a school attended only by students who would be excluded from taking the assessment for intellectual, functional, or linguistic reasons was considered a school-level exclusion).

Within a selected in-scope school, the following categories of teachers were excluded from the sample:

- teachers teaching in schools exclusively serving special needs students
- teachers who also act as school principals: no teacher data collected, but school principal data collected
- substitute, emergency or occasional teachers
- teachers on long-term leave.

Within a selected in-scope school, the following categories of students were excluded from the sample:

- 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 (i.e. students who 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; these 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.

However, 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.

Sample size requirements

In each country/economy, a sample of 150 schools (unless discussions with the national project manager led to a different size) was drawn randomly from the sample of schools drawn for PISA 2018. Within each participating school, 35 students were randomly selected. However, in schools with 35 or fewer eligible students, all students were selected. As the PISA 2018 data collection proceeded, the set of (original sample or replacement) schools participating in PISA emerged and, thus, the set of schools where the TALIS-PISA link 2018 should be administered. Within each of the schools that had participated in PISA 2018 and also sampled for the TALIS-PISA link 2018, the school principal and a random sample of 20 teachers teaching 15-year-old students were surveyed. However, in schools with 20 or fewer eligible teachers, all teachers were selected. Thus, the nominal international sample size of the TALIS-PISA link 2018 was set at 150 schools, 3 000 teachers and 5 250 students for each participating country or economy.

Weights

Final weights allow the production of country-level estimates from the observed sample data. The estimation weight indicates how many population units are represented by a sampled unit. The final weight is the combination of many factors, reflecting the probabilities of selection at the various stages of sampling and the response obtained at each stage. To maintain the unbiasedness of the estimates, other factors may also come into play as dictated by special conditions (e.g. adjustment for teachers working in more than one school).

Estimating the sampling error for surveys with complex designs, such as TALIS and PISA, requires special attention. TALIS and PISA adopted the balanced repeated replication (BRR) for estimation of the sampling error of the estimates. BRR is a replication method suited to sample designs where exactly two primary sampling units (PSUs) are selected in each stratum. It leads to (approximately) unbiased estimates of sampling error.

As mentioned above, the sample of schools for the TALIS-PISA link was a subset of the sample of schools selected to take part in PISA 2018. Given the sequencing of events between TALIS and PISA, the sampling team could not limit sub-sampling for the TALIS-PISA link to schools that had participated in PISA. Therefore, they had to draw the sub-sample from the full sample of schools prior to the PISA data collection. However, because data collection for the TALIS-PISA link was scheduled to take place after completion of the data collection for PISA (at least, for any given school), the school base weight was that of the PISA 2018 design, adjusted for sub-sampling.

The final TALIS-PISA link school weight (estimation weight [SCHWGT]) is the product of the TALIS-PISA link school base weight and the TALIS-PISA link school non-response adjustment factor. The final TALIS-PISA link teacher weight (estimation weight [TCHWGT]) is the product of the TALIS-PISA link teacher base weight, the three adjustment factors associated with each participating teacher (i.e. non-response adjustment within the school, and multiplicity and exclusion adjustments) and the final TALIS-PISA link school weight. Balanced repeated replicate weights for both teacher and school observations [TRWGT1-TRWGT100 and SRWGT1-SRWGT100] can be used to obtain (approximately) unbiased estimates of sampling errors.

To conduct student-level analyses based on the student-level merged TALIS-PISA dataset (i.e. student data merged with principal data and teacher data aggregated at the school level), the final TALIS-PISA link student weight (estimation weight) as well as the TALIS-PISA link student-level balanced repeated replicate weights need to be estimated. These can be done using the following steps:

- Compute the within-school component of student weights (i.e. the student base weight, adjusted for non-response, and trimmed if needed – for more detail, see Chapter 8 of the PISA 2018 Technical Report (OECD, 2020[2]) – by dividing the final student weight included in the PISA 2018 student dataset [W_FSTUWT] by the final school weight included in the PISA 2018 school dataset [W_SCHGRNRABWT].[2]
- Compute the final TALIS-PISA link student weight as the product of the within-school component of student weights and the final TALIS-PISA link school weight included in the TALIS-PISA link 2018 school/principal dataset [SCHWGT]. This ensures that the non-response adjustments accounting for school non-response specific to the TALIS-PISA link sub-sample are applied.[3]
- Compute the TALIS-PISA link student-level balanced repeated replicate weights by multiplying the within-school component of student weights by each of the 100 school-level balanced repeated replicate weights that are included in the TALIS-PISA link school/principal dataset [SRWGT1-SRWGT100].

Response rate requirements

The technical standards for both TALIS-PISA link 2018 and PISA 2018 outlined the response rate (participation rate) requirements for their respective target populations (i.e. schools, teachers and students). Although reaching the required levels of participation does not preclude some degree of error in the results, it should reduce reliance on the "missing at random" assumptions made for the non-response weighting adjustments.

TALIS-PISA link 2018 set the minimum school response rate at 75% of sampled eligible and non-excluded schools after replacement. TALIS-PISA link considered schools where the principal returned a questionnaire to be "participating" schools for the purposes of the school weights and database. Although replacement schools could be called upon as substitutes for non-responding schools, the study's national project managers were encouraged to do all they could to obtain the participation of the schools in the original sample. Countries that reached less than 75% school participation after replacement had to demonstrate convincingly that their sample was not significantly biased. The minimum response rate for teachers was set at 75% of all sampled teachers across all participating schools. TALIS-PISA link considered schools where at least 50% of selected teachers responded to the questionnaire to be "participating" schools for the purposes of the teacher weights and database, regardless of their participation status on the school database, that is, regardless of whether or not their principal returned his or her questionnaire.

PISA 2018 set the minimum school response rate at 85% of sampled eligible and non-excluded schools. If the initial school response rate fell between 65% and 85%, an acceptable school response rate could still be reached through the use of replacement schools. Furthermore, a school with a student participation rate between 25% and 50% was not considered as a participating school for the purposes of calculating and documenting response rates. However, data from such schools were included in the database and contributed to the estimates included in the initial PISA international report. Data from schools with a student participation rate of less than 25% were not included in the database, and such schools were regarded as non-respondents. The minimum response rate for students was set at 80% of all sampled students across all participating schools. A minimum student response rate of 50% within each school was required for a school to be regarded as participating: the overall student response rate was computed using only students from schools with at least a 50% student response rate.

Adjudication process

Data adjudication is the process through which each national dataset is reviewed and a judgement about the appropriateness of the data for the main reporting goals is formed. For both TALIS-PISA link 2018 and PISA 2018, the basic principle that guided the adjudication was to determine, for each participating country and economy, whether the data released to the countries and economies are fit to provide policy relevant, robust international indicators and analysis on students, teachers and teaching in a timely and cost effective manner.

To establish fitness for use, a number of quality assurance processes were designed and activated throughout the survey process. Some processes relied on expert advice and opinion; some relied on qualitative information and learned judgement; some relied on quantitative information. School, teacher and student data received separate adjudication evaluation per country and economy. The issues evaluated concerned the questionnaire adaptation to national context, translation and verification, quality of the sampling frame, handling of out-of-scope and refusal units, within-school sampling, data collection, data cleaning, the reports of quality observers, participation rates and overall compliance with the technical standards. Once each survey process had been assessed, a recommended rating was formulated, accounting for the participation rates, and for any unresolved issue.

The adjudication of the TALIS-PISA link 2018 samples had to wait until the PISA 2018 samples had been adjudicated, as the former was dependent on the latter to allow the final determination of the recommended rating. Even if the recommended rating for TALIS-PISA link, based solely on what happened during the preparation and collection of the TALIS-PISA link, was "good", if the data or samples from PISA were to be rated less favourably, the matched file could not be adjudicated as "good". It could only be adjudicated as the weakest, at most, of either rating. For more detailed information, please refer to the TALIS 2018 Technical Report (OECD, 2019[1]) and the PISA 2018 Technical Report (OECD, 2020[2]).

The adjudication rules for the TALIS-PISA link 2018 samples, based on participation rates for principals and teachers, are displayed in Tables A A.1 and A A.2. An explanation of the codes used is given below.

Table A A.1. Adjudication rules for school or principal data in TALIS 2018

School participation (returned principal questionnaires)		Risk of school non-response bias	Rating
Before replacement	After replacement		
≥75%	≥75%		Good
50% - 75%	≥75%		Fair (A)
	50% - 75%	Low	Fair (C)
		High	Poor (D)
<50%			Insufficient

Table A A.2. Adjudication rules for teacher data in TALIS 2018

School participation (minimum teacher participation)		Teacher participation after school replacement	Risk of teacher non-response bias	Rating
Before replacement	After replacement			
≥75%	≥75%	≥75%		Good
		50% - 75%		Fair (A)
50% - 75%	≥75%	≥75%		Fair (B)
		50% - 75%	Low	Fair (C)
			High	Poor (D)
50% - 75%	50% - 75%			Poor (E)
< 50%	≥75%			Poor (F)
< 50%	< 75%			Insufficient

The following bulleted list is a simple guide aimed at helping data users appreciate the limitations on use or quality:

- **Good:** the participating country's/economy's data can be used for all reporting and analytical purposes and can be included in international comparisons.
- **Fair (A):** national and sub-national estimates can be produced; some teacher characteristics may suffer from a larger standard error (SE), hence the warning "Fair" and no additional warnings to users appear necessary.
- **Fair (B, only for teacher data adjudication):** national and sub-national estimates can be produced; some sub-national estimates may be of lower precision (larger SE) if sample size is locally low, hence the warning "Fair" and no additional warnings to users appear necessary.
- **Fair (C):**
  - national and sub-national estimates can be produced.
  - some sub-national estimates may be of lower precision (larger SE) if sample size is locally low, hence the warning "Fair", but a note on data quality could appear pointing to the outcome of the non-response bias analysis (NRBA).

  - since school participation is somewhat lower than under (B), comparing sub-national estimates should be done with care, as some of those results are based on few schools.
  - comparing small sub-national estimates with similar groups from other countries is likely to uncover any statistically meaningful differences, as the SE are likely too large.
- **Poor (D):**
  - in addition to the warnings issued for the previous category, a note should warn users of indications of non-response biases in some estimates.
  - comparisons of sub-national estimates should be limited to the groups with the larger sample sizes.
  - at this point, the sample represents between 37% and 56% of the teaching workforce, from a rather small sample of schools.
  - comparisons with similar groups in foreign countries would not be encouraged.
- **Poor (E, only for teacher data adjudication):** sub-national estimates would not be recommended; there should be a note pointing out the difficulty of obtaining a representative sample of schools.
- **Poor (F, only for teacher data adjudication):** limitations similar to those of line E, but there should be a note pointing out the difficulty of obtaining at least 50% participation of the selected sample of schools; risks of having a non-representative sample of schools.
- **Insufficient:** weights should not be calculated for any official tabulations; hence, data should not be incorporated into international tables, models, averages, etc.

Tables A A.3 and A A.4 display the participation rates for the principals and teachers in each country/economy that participated in the TALIS-PISA link.

Table A A.3. TALIS-PISA link 2018: Principal’s participation and recommended ratings

Participating country/economy	Number of schools sub-sampled	Number of eligible schools	Number of participating principals	Principals’ participation before replacement (%)	Principals’ participation after replacement (%)	Recommended rating
Australia	150	148	131	66.9	88.5	Fair
Ciudad Autónoma de Buenos Aires (Argentina)	104	81	77	88.9	95.1	Good
Colombia	162	162	153	91.4	94.4	Good
Czech Republic	190	182	175	96.8	96.8	Good
Denmark	150	150	83	52.0	57.8	Poor
Georgia	150	144	124	86.1	86.1	Good
Malta	63	50	47	94.0	94.0	Good
Turkey	150	147	142	96.6	96.6	Good
Viet Nam	150	115	115	100.0	100.0	Good

Note: There are five schools – three in Colombia and two in Georgia – that cannot be linked to PISA data. These schools were originally sampled for PISA, but were finally left out of the final PISA database. However, they were sampled for the TALIS-PISA link and participated in it, so they were included in the final TALIS-PISA link datasets.

Table A A.4. TALIS-PISA link 2018: Teacher's participation and recommended ratings

Participating country/ economy	School sample size	Number of eligible schools	Number of participating schools	Number of participating teachers	Estimated size of teacher population	School participation before replacement (%)	School participation after replacement (%)	Teachers' participation in participating schools (%)	Overall teacher participation (%)	Recommended rating
Australia	150	148	131	2 233	34 598	65.6	88.8	93.4	82.9	Good
CABA (Argentina)[1]	104	81	73	1 194	2 673	87.7	90.1	85.1	76.7	Good
Colombia	162	162	154	2 242	179 900	91.4	95.1	94.6	89.9	Good
Czech Republic	190	182	173	2 592	62 040	95.1	95.1	94.8	90.0	Good
Denmark	150	150	100	1 079	20 777	65.8	70.0	85.9	60.2	Poor
Georgia	150	144	132	1 923	24 592	93.1	93.1	94.3	87.8	Good
Malta	63	50	44	857	1 102	88.0	88.0	88.6	78.0	Good
Turkey	150	147	142	3 591	236 904	97.9	97.9	99.6	97.5	Good
Viet Nam	150	115	114	2 170	250 645	99.3	99.3	98.4	97.7	Good

1. Ciudad Autónoma de Buenos Aires (Argentina)
Note: There are five schools – three in Colombia and two in Georgia – that cannot be linked to PISA data. These schools were originally sampled for PISA, but were finally left out of the final PISA database. However, they were sampled for the TALIS-PISA link and participated in it, so they were included in the final TALIS-PISA link datasets.

Merging the TALIS-PISA link 2018 and PISA 2018 datasets

Data from participating school principals, participating teachers and participating students from the same schools can be merged as result of the school-level link.

Table A A.5 displays the number of observations for each category of participants, once their data are merged at the school level. The school-level merged TALIS-PISA dataset (i.e. student data aggregated at the school level merged with principal data and teacher data aggregated at the school level) has 1 058 school-level observations. The teacher-level merged TALIS-PISA dataset (i.e. teacher data merged with principal data and student data aggregated at the school level) has 17 809 teacher-level observations. The student-level merged TALIS-PISA dataset (i.e. student data merged with principal data and teacher data aggregated at the school level) has 31 077 student-level observations.

Table A A.5. Number of observations in the merged TALIS-PISA link 2018 datasets

Participating country/economy	Number of participating schools/principals	Number of participating teachers	Number of participating students
Australia	131	2 233	2 404
Ciudad Autónoma de Buenos Aires (Argentina)	73	1 194	2 111
Colombia	151	2 199	4 699
Czech Republic	173	2 592	3 992
Denmark	100	1 079	2 269
Georgia	130	1 894	3 064
Malta	44	857	2 988
Turkey	142	3 591	5 528
Viet Nam	114	2 170	4 022
Total	1 058	17 809	31 077

References

OECD (2020), *PISA 2018 Technical Report*, https://www.oecd.org/pisa/data/pisa2018technicalreport/ (accessed on 16 November 2020). [2]

OECD (2019), *TALIS 2018 Technical Report*, OECD, Paris, http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf. [1]

Notes

[1] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[2] Non-response adjustment applied to the within-school component of student weights is, in some cases, based on “non-response classes” from the full PISA dataset (including schools that are not part of the TALIS-PISA link sub-sample). However, this is considered to be a minor issue.

[3] Since the TALIS-PISA link dataset is a subset of the PISA dataset, school non-response adjustments specific to the TALIS-PISA link sub-sample also account for non-participation in PISA. However, there are five schools – three in Colombia and two in Georgia – that participated in TALIS-PISA link 2018, but were considered non-participants in PISA 2018. These schools were originally sampled for PISA, but were finally left out of the final PISA database. However, they were sampled for the TALIS-PISA link and participated in it so they were included in the final TALIS-PISA link datasets.

Annex B. Technical notes on analyses in this report

Variables used in the analyses

This report explores the teacher and school characteristics and practices that matter for student outcomes. The outcome indicators included in the analyses are sourced from the Programme for International Student Assessment (PISA) 2018 student data. These student outcomes are:

- Student achievement in the PISA score in reading, mathematics and science (Chapters 2 and 4).
- Student social and emotional outcomes, such as the indices of classroom disciplinary climate, teacher enthusiasm and student perception of difficulty of the PISA test, as well as the dummy variable of student expectation of completing at least a tertiary degree (Chapter 3).
- School-level disparities in outcomes between girls and boys that are defined as the difference between the average school-level PISA score for girls minus the average school-level PISA score for boys (Chapter 4). Differences are positive when they are in favour of girls and negative when they are in favour of boys.

In order to make the most of the TALIS-PISA link,[1] a broad range of variables were included in the analyses aiming to identify the teacher and school characteristics and practices that matter for student outcomes. The teacher and school factors included in the analyses represent variables from the TALIS components (both teacher and principal datasets) of the TALIS-PISA link 2018 data. The selection of teacher and school variables was guided by theory and previous research findings. Table A B.1 provides an overview of the 18 teacher and school dimensions and the almost 150 variables within these dimensions that are included in the analyses. These variables are jointly introduced in lasso regressions (see below) and are introduced separately by blocks for each dimension in standard regression analyses (see below).

However, apart from relying on theory and previous research findings, other considerations were also taken into account to guide the initial selection of TALIS indicators used for the analyses. Notably, it was deemed important to:

- Limit the loss of observations, due to missing values, given the large number of predictors included in the analyses. This was ensured mainly by opting for the inclusion of teacher questionnaire variables instead of principal questionnaire variables whenever it was possible. Indeed, because teacher data are aggregated at the school level, variables derived from the teacher questionnaire are less likely to have missing values.
- Include only those variables that were administered in all TALIS-PISA link participating countries and economies.
- Prioritise the use of complex scales that were available in the TALIS-PISA link datasets over the individual items contributing to these scales.

Due to the survey design of the TALIS-PISA link, teachers and students can be linked only at the school and not at the classroom level. In other words, the data do not allow matching teachers with their students; rather, the data only allow matching a sample of teachers teaching 15-year-old students in a school with a sample of 15-year-old students of that same school. Therefore, information that is based on teachers' responses is always averaged at the school level within this report. Depending on the analysis, variables based on teachers' responses are averaged either for all teachers within the school, or only for subject domain teachers (i.e. reading, mathematics or science teachers) within the school. For more detail on the share of schools sampled within the TALIS-PISA link by the number of subject domain teachers, see Table A B.2.

Most of the analyses presented in this report also include controls for student characteristics, such as student gender, migrant background and socio-economic status, based on data available from the PISA student dataset. In the case of the school-level analyses of Chapter 4, these student characteristics are averaged at the school level and introduced as controls.

Lasso regression analysis

The least absolute shrinkage and selection operator (also known as lasso), which is a machine learning technique within the family of supervised statistical learning methods, is applied throughout this report as a compass to guide the selection of key teacher and school factors related to student achievement, social-emotional skills and gaps in student performance within schools. In this report, lasso is used for model selection, although it can be used for prediction and inference as well. Lasso has several attributes that makes it an attractive tool for selecting among the many variables collected through the TALIS questionnaires those that are potentially key predictors of student outcomes. These attributes are:

- Lasso is designed to select variables that are important and should be included in the model.
- The outcome variable guides the model selection process (i.e. supervised statistical learning method).
- Lasso can handle high-dimensional models where the number of variables is high relative to the number of observations.

Lasso is most useful when only a few out of many potential variables affect the outcome (Hastie, Tibshirani and Friedman, 2017[1]; Hastie, Tibshirani and Wainwright, 2015[2]; Tibshirani, 1996[3]). The assumption that the number of coefficients that are non-zero (i.e. correlated with the outcome variable) in the true model is small relative to the sample size is known as a sparsity assumption. The approximate sparsity assumption requires that the number of non-zero coefficients in the model that best approximates the true model be small relative to the sample size.

Lasso estimates coefficients in a model. It selects variables that correlate well with the outcome in one dataset (training sample) and then tests whether the selected variables predict the outcome well in another dataset (validation sample). Lasso proceeds with model selection by estimating model coefficients in such a way that some of the coefficient estimates are exactly zero and, hence, excluded from the model, while others are not (Hastie, Tibshirani and Friedman, 2017[1]; Hastie, Tibshirani and Wainwright, 2015[2]; Tibshirani, 1996[3]). In the context of model selection, lasso may not always be able to distinguish an irrelevant predictor that is highly correlated with the predictors in the true model from the true predictors (Wang et al., 2019[4]; Zhao and Yu, 2006[5]).

Lasso for linear models solves an optimisation problem. The lasso estimate is defined as:

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} \omega_j |\beta_j| \right\}$$

Where y is the outcome variable, x refers to the potential covariates, β is the vector of coefficients on X, λ is the lasso penalty parameter, ω refers to the parameter-level weights known as penalty loadings and $\sum_1^p |\beta|$ is the L_1 lasso penalty. As the lasso penalty term is not scale invariant, one needs to standardise the variables included in the model before solving the optimisation problem.

Thus, the optimisation problem contains two parts, the least-squares fit measure:

$$\frac{1}{2N}\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2$$

And the penalty term:

$$\lambda\sum_{j=1}^{p}\omega_j|\beta_j|$$

The λ and ω parameters (also called "tuning" parameters) specify the weight applied to the penalty term. When λ is large, the penalty term is also large, which results in lasso selecting few or no variables. As λ decreases, the penalty associated with each non-zero β decreases, which results in an increase in the number of coefficient estimates kept by lasso. When $\lambda = 0$, then lasso reduces to the ordinary least squares (OLS) estimator without any coefficient estimates being excluded from the model.

Two commonly used methods to select the so called "tuning" parameters are cross-validation (CV), and the adaptive lasso. CV finds the λ^* that minimises the out-of-sample prediction error. Although CV works well for prediction, it tends to include covariates whose coefficients are zero in the true model that best approximates the data. The adaptive lasso, which consists of two CVs, is more parsimonious when it comes to model selection. After finding a CV solution for λ^*, it does another CV among the covariates selected in the first step by using weights ($\omega = 1/|\hat{\beta}|$, where $\hat{\beta}$ are the penalised estimates from the first CV) on the coefficients in the penalty function. Covariates with smaller coefficients are more likely to be excluded in the second step (Drukker and Lui, 2019[6]).

The third commonly used method is the plugin lasso. Among the three methods, the plugin lasso is the most parsimonious and also the fastest one in terms of computational time. Instead of minimising a CV function as presented above for the CV and adaptive lasso methods, the plugin function uses an iterative formula to find the smallest value of λ that is large enough to dominate the estimation error in the coefficients. The plugin lasso selects the penalty loadings to normalise the scores of the (unpenalised) fit measure for each parameter and then it chooses a value for λ that is greater than the largest normalised score with a probability that is close to 1 (Drukker and Lui, 2019[6]). For more detail on the plugin lasso, see Belloni et al. (2012[7]) and Drukker and Liu (2019[8]). The plugin lasso tends to select the most important variables and it is good at not including covariates that do not belong to the true model. However, unlike the adaptive method, the plugin lasso can overlook some covariates with large coefficients and select covariates with small coefficients (Drukker and Lui, 2019[6]). Given its favourable model selection attributes, the plugin method is applied in Chapters 2 and 3.

Yet, in the school-level analysis of Chapter 4, which explores the teacher and school factors that could play a role in mitigating within-school disparities in performance between girls and boys, the adaptive lasso is used. Due to the analysis being conducted at the school level rather than at the student level, sample sizes decrease considerably, leading to the selection of fewer variables by lasso. Therefore, the use of adaptive lasso is preferred, as it results in more covariates being selected as compared to the more parsimonious plugin lasso. It is also important to note that the model selection properties of lasso have limitations. Notably, irrespective of the way in which the tuning parameters are selected, lasso may not always be able to distinguish an irrelevant predictor that is highly correlated with the predictors in the true model from the true predictors.

Dividing the sample into training and validation sub-samples allows for validating the performance of the lasso estimator (or estimators, if different methods to select the tuning parameters are tested).[2] In this report, the training and validation samples are generated by randomly splitting the overall TALIS-PISA link sample into two sub-samples, with 85% of the observations allocated to the training sample and 15% kept for the validation sample. While the proportion of observations allocated to each sub-sample could be considered somewhat arbitrary, sensitivity analysis shows that the lasso regression results reported herein are fairly robust to how the TALIS-PISA link sample is split into training and validation sub-samples.[3] It is also important to note that sample split is performed after creating a balanced sample that includes only those observations that have full information (i.e. observations with missing information are excluded) for all variables included in the model.[4]

Applying lasso for model selection means finding a model that fits the data, not finding a model that allows for interpreting estimated coefficients as effects. Thus, when used for model selection, lasso selects variables and estimates coefficients, but it does not provide the standard errors required for performing statistical inference. Indeed, lasso's covariate-selection ability makes it a non-standard estimator and prevents the estimation of standard errors.

In this report, lasso is applied for model selection based on the overall population of 15-year-old students (Chapters 2 and 3) and schools (Chapter 4) surveyed within the TALIS-PISA link (i.e. the pooled sample across all participating countries and economies). As the model selection is based on the pooled sample, country fixed effects are imposed on lasso to ensure they are always included among the selected covariates. In addition, the controls for student characteristics – such as student gender, migrant background and socio-economic status – are also imposed on lasso to ensure they are always included among the selected covariates. Moreover, sampling weights are not used in the lasso regression analysis. This may be a limitation; hence, caution is warranted while interpreting the results of lasso regression analyses within this report.

Lasso regressions are estimated using the Stata (version 16.1) (StataCorp, 2019[9]) "lasso" module – see Drukker and Lui (2019[6]) for an introduction.

Standard regression analysis

Standard regression analyses are conducted to examine the teacher and school dimensions that explain most of the differences in school average performances (i.e. variance decomposition analysis) and also to explore the country-level relationships between student outcomes, or within-school gender gaps in student achievement (in the case of Chapter 4), and teacher and school dimension (taken separately). In comparison to lasso, standard regressions provide the confidence intervals of the coefficient estimates that, in turn, allow for drawing inferences about the overall population. Moreover, they lead to more accurate coefficient estimates through the introduction of final and balanced repeated replicate weights and the use of plausible values of student performance. Multiple linear regression is used in those cases where the dependent (or outcome) variable is considered continuous. Binary logistic regression is employed when the dependent (or outcome) variable is a binary categorical variable. Regression analyses are carried out for each country separately. The TALIS-PISA link average refers to the arithmetic mean of country-level estimates.

Control variables are included in the standard regression models, with the exception of the variance decomposition analyses. The control variables are selected based on theoretical reasoning and, preferably, limited to the most objective measures or those that do not change over time. Controls for student characteristics include student's gender, migrant background and socio-economic status. Controls for the average classmates' characteristics within the school include: the share of students whose first language is different from the language(s) of instruction, low academic achievers, students with special needs, students with behavioural problems, students from socio-economically disadvantaged homes, academically gifted students, students who are immigrants or with a migrant background and students who are refugees. Controls for classmates' characteristics are only included in the analyses presented in Chapters 2 and 3. Controls for classmates' characteristics are excluded from the analyses featured in Chapter 4.[5]

The controls for the characteristics of students and classmates are introduced into the models in steps. This approach also requires that the models at each step be based on the same sample. Each regression model for each teacher or school dimension is estimated based on the same restricted and balanced sample with full information (i.e. observations with missing information are excluded) for all variables included in the analyses. The sample only varies according to how teacher variables are averaged at the school level, i.e. if the focus is on all teachers or teachers of a given subject.

Multiple linear regression analysis

Multiple linear regression analysis provides insights into how the value of the continuous outcome variable changes when any one of the explanatory variables varies while all other explanatory variables are held constant. In general, and with everything else held constant, a one-unit increase in the explanatory variable (x_i) increases, on average, the outcome variable (Y) by the units represented by the regression coefficient (β_i):

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i + \varepsilon$$

When interpreting multiple regression coefficients, it is important to keep in mind that each coefficient is influenced by the other explanatory variables in a regression model. The influence depends on the extent to which explanatory variables are correlated. Therefore, each regression coefficient does not capture the total effect of explanatory variables on the outcome variable. Rather, each coefficient represents the additional effect of adding that variable to the model, considering that the effects of all other variables in the model are already accounted for. It is also important to note that, because cross-sectional survey data are used in these analyses, no causal conclusions can be drawn.

Regression coefficients in bold in the data tables presenting the results of regression analysis (included in Annex C) are statistically significantly different from 0 at the 95% confidence level.

Binary logistic regression analysis

Binary logistic regression analysis enables the estimation of the relationship between one or more explanatory variables and an outcome variable with two categories. The regression coefficient (β) of a logistic regression is the estimated increase in the log odds of the outcome per unit increase in the value of the predictor variable.

More formally, let Y be the binary outcome variable indicating no/yes with 0/1, and p be the probability of Y to be 1, so that $p = prob(Y = 1)$. Let $x_1, \ldots x_k$ be a set of explanatory variables. Then, the logistic regression of Y on $x_1, \ldots x_k$ estimates parameter values for $\beta_0, \beta_1, \ldots, \beta_k$ via the maximum likelihood method of the following equation:

$$Logit(p) = \log(p/(1-p)) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

Variance decomposition analysis

Variance decomposition analysis is applied to complement the findings from the lasso regressions as it can reveal the relative importance of each teacher and school dimension in explaining the average differences in student performances across schools (and ultimately, variance in student achievement). The share of between-school variance explained by a teacher or school dimension j is estimated as the ratio of the variance explained by dimension j ($R^2_{Dim.j}$) to the total variance in student outcomes explained at the school level (R^2_{Total}), hence:

$$Between - school\ VAR_{Dim.j} = R^2_{Dim.j}/R^2_{Total}$$

where the total variance in student outcomes explained at the school level (R^2_{Total}) is estimated as the R^2 of a linear regression model with the outcome variable (Y) regressed on school fixed effects. R^2 represents the proportion of the observed variation in the outcome variable that can be explained by the explanatory variables. Similarly, the share of variance explained by dimension j ($R^2_{Dim.j}$) is estimated as the R² of a linear regression model with the outcome variable (Y) regressed on the variables included in dimension j.

Chapter 3 of this report features a binary outcome variable, students' educational expectations, as a measure of student interest in school. PISA measures educational expectations by asking students which educational level they expect to complete. Their responses are used to create a dummy variable that equals 1 if the student expects to complete at least a tertiary degree and, otherwise, equals 0. Since student educational expectation is coded as a binary variable, the share of variance cannot be estimated as it can be for the continuous variables. Indeed, binary logistic regressions cannot provide a goodness-of-fit measure that would be equivalent to the R². Unlike linear regressions with normally distributed residuals, it is not possible to find a closed-form expression for the coefficient values that maximise the likelihood function of logistic regressions; thus, an iterative process must be used instead. Yet, the goodness-of-fit of binary logistic models can be evaluated by the pseudo-R².[6] Similarly to the R², the pseudo-R² also ranges from 0 to 1, with higher values indicating better model fit. Nevertheless, pseudo-R² cannot be interpreted as one would interpret the R². In Chapter 3, the pseudo-R² of the logistic regression of student educational expectation on school fixed effects is used as a proxy for the percentage of total variance in student expectation of completing at least a tertiary degree, explained at the school level. Then, the share of variance explained by dimension j ($R^2_{Dim.j}$) is approximated with the pseudo-R² of a logistic regression model with the outcome variable (Y) regressed on the variables included in dimension j.

It has to be noted that the variance decomposition analysis presented within this report has a limitation, as the shares of between-school variance explained by each dimension may be artificially driven by the number of variables included in a given dimension. Indeed, the dimensions that have the lowest explanatory power tend to include few variables, while the number of variables included in the dimensions that explain the largest shares of the differences in school average performance is high in comparison to other dimensions. Thus, caution is warranted when interpreting these results.

Quantile regressions

In contrast with standard linear regression, which estimates the conditional mean of the outcome variable given a set of explanatory variables, quantile regression provides information about the association between the outcome variable and the explanatory variables at the different points in the conditional distribution of the outcome variable (Koenker, 2017[10]; Koenker, 2005[11]; Koenker and Bassett, 1978[12]). Quantile regression estimates an equation expressing a quantile (or percentile) of the conditional distribution of the outcome variable as a linear function of the explanatory variables. At the qth quantile, the quantile regression estimator, $\widehat{\beta_q}$, minimises over β_q the objective function (Cameron and Trivedi, 2009[13]):

$$Q(\beta_q) = \sum_{i:y_i \geq x_i'\beta}^{N} q|y_i - x_i'\beta_q| + \sum_{i:y_i < x_i'\beta}^{N} (1-q)|y_i - x_i'\beta_q|$$

Where $0 < q < 1$ and different choices of q estimate different values of β. The higher the value of q, the more weight is placed on prediction for observations with $y \geq x'\beta$ than for observations with $y < x'\beta$. The objective function is optimised using linear programming methods. The estimator that minimises $Q(\beta_q)$ has well-established asymptotic properties.

Quantile regression has several attributes that make its use attractive as compared to standard linear regression (Cameron and Trivedi, 2009[13]). For instance, it provides a richer characterisation of the relationship between the outcome variable and the explanatory variables by allowing the effects of the explanatory variables to vary over different quantiles of the conditional distribution. Notably, it allows for examining the impact of explanatory variables on both the location and scale parameters of the model. Moreover, quantile regression is more robust to outliers and also in terms of the assumptions about the distribution of regression errors. Namely, it does not require assumptions about the parametric distribution of regression errors. Hence, quantile regression is a suitable tool to analyse models characterised by a change in the variance of the error terms (i.e. heteroskedasticity).

However, it is important to note that quantile regression estimates tend to be more precise at the centre of the distribution as compared to upper and lower quantiles (meaning that standard errors tend to be smaller at the centre of the distribution) (Cameron and Trivedi, 2009[13]). Thus, for those relationships where the variation in the effects of the explanatory variables over the different quantiles of the conditional distribution of the outcome variable is limited, quantile regression tends to be less likely to find significant regression coefficients at the tails of a distribution than at the centre.

Use of student, teacher and school weights

The TALIS-PISA link and PISA samples were collected following a stratified two-stage probability sampling design. This means that teachers and students (second stage units, or secondary sampling units) were to be randomly selected from the list of in-scope teachers and students in each of the randomly selected schools (first stage units, or primary sampling units). For these statistics to be meaningful for a country, they need to reflect the whole population from which they were drawn and not merely the sample used to collect them. Thus, survey weights must be used in order to obtain design-unbiased estimates of population or model parameters. Except for the lasso regression analysis, survey weights are used in all other analyses presented in this report.

The analyses presented in Chapters 2 and 3, as well as the quantile regression analysis included in Chapter 4, are based on the student-level merged TALIS-PISA dataset (i.e. student data merged with principal data and teacher data aggregated at the school level). The statistics resulting from the standard regression analyses are estimated using the final TALIS-PISA link student weight (estimation weight), as well as the TALIS-PISA link student-level balanced repeated replicate weights (for more detail, see Annex A).

In Chapter 4, the analysis conducted at the school level that explores the teacher and school factors that could play a role in mitigating within-school disparities in performance between girls and boys is based on the school-level merged TALIS-PISA dataset (i.e. student data aggregated at the school level merged with principal data and teacher data aggregated at the school level). The statistics resulting from the standard regression analysis are estimated using final TALIS-PISA link school weight (estimation weight), as well as the TALIS-PISA link school-level balanced repeated replicate weights (for more detail, see Annex A).

The teacher-level merged TALIS-PISA dataset (i.e. teacher data merged with principal data and student data aggregated at the school level) is not used for this report. Nevertheless, the final TALIS-PISA link teacher weights are used for averaging teachers' responses at the school level.

Standard errors and significance tests

The statistics in this report represent estimates based on samples of teachers and principals, rather than values that could be calculated if every teacher and principal in every country had answered every question. Consequently, it is important to measure the degree of uncertainty of the estimates. Hence, each estimate presented in this report, with the exception of lasso regression results, has an associated degree of uncertainty that is expressed through a standard sampling error. When used for model selection, lasso selects variables and estimates coefficients, but it does not provide the standard errors required for performing statistical inference. Yet, standard errors are computed and presented for the estimates of all other analyses. The use of confidence intervals provides a way to make inferences about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. From an observed sample statistic and assuming a normal distribution, it can be inferred that the corresponding population result would lie within the confidence interval in 95 out of 100 replications of the measurement on different samples drawn from the same population. The reported standard errors were computed with a balanced repeated replication (BRR) methodology.

The analyses within this report that focus on student achievement (Chapters 2 and 4) are based on plausible values. PISA report student performance through plausible values in order to account for measurement error (OECD, 2009[14]). This error results from the fact that no test can perfectly measure proficiency in broad subjects, such as reading, mathematics and science. Hence, plausible values can be considered as a representation of the range of abilities that a student might reasonably have.[7] In turn, the standard errors reported for statistics based on plausible values do not only include sampling error but also measurement error.

International averages

The TALIS-PISA link average corresponds to the arithmetic mean of the respective country estimates with available data. In Chapter 2, the TALIS-PISA link average covers all participating countries and economies, excluding Viet Nam.[8] In Chapter 3, the TALIS-PISA link average covers all participating countries and economies. In Chapter 4, the TALIS-PISA link average excludes Viet Nam in the case of student-level analyses and it excludes Malta[9] and Viet Nam in the case of school-level analyses.

In the case of some countries, data may not be available for specific indicators, or specific categories may not apply. Therefore, readers should keep in mind that the term “TALIS-PISA link average” refers to the countries included in the respective averages. Each of these averages may not necessarily be consistent across all columns of a table.

Robustness of TALIS-PISA link data to grade repetition

Grade repetition can potentially introduce bias into the estimates of school performance, estimated as the student performance averaged at the school level. Indeed, if grade repetition is common and grade repeaters tend to attend different grade levels than their peers who did not repeat a grade, and if those different grade levels belong to different schools, then estimates of school performance can be biased. Indeed, school performance would be overestimated in schools attended by non-grade repeaters and underestimated in schools attended by grade repeaters. Among the countries and economies participating in the TALIS-PISA link, the retention rate seems especially high in Ciudad Autónoma de Buenos Aires (Argentina) (henceforth CABA [Argentina]) and Colombia, where more than 75% of students enrolled at ISCED level 2 and fewer than 15% of students enrolled at ISCED level 3 are grade repeaters (Table A B.3). However, these two countries/economies also happen to be the ones with the highest share of sampled schools (greater than 80%) that enrol students at both ISCED levels 2 and 3 (Table A B.4). Hence, students who have repeated a grade are often enrolled in the same school as students who have not. Thus, students sampled in a school can be considered as being representative of the school, even in CABA (Argentina) and Colombia.

References

Belloni, A. et al. (2012), “Sparse models and methods for optimal instruments with an application to eminent domain”, *Econometrica: Journal of the Econometric Society*, Vol. 80/6, pp. 2369-2429, http://dx.doi.org/10.3982/ecta9626. [7]

Cameron, A. and P. Trivedi (2009), *Microeconometrics Using Stata*, Stata Press, College Station, TX. [13]

Drukker, D. and D. Liu (2019), *A plug-in for Poisson lasso and a comparison of partialing-out Poisson estimators that use different methods for selecting the lasso tuning parameters*. [8]

Drukker, D. and D. Lui (2019), “An introduction to the lasso in Stata”, *The Stata Blog*, https://blog.stata.com/2019/09/09/an-introduction-to-the-lasso-in-stata/. [6]

Hastie, T., R. Tibshirani and J. Friedman (2017), “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”, in *Springer Series in Statistics*, Springer, New York, https://web.stanford.edu/~hastie/ElemStatLearn//printings/ESLII_print12.pdf. [1]

Hastie, T., R. Tibshirani and M. Wainwright (2015), “Statistical learning with sparsity: The lasso and generalizations”*, Monographs on Statistics and Applied Probability*, No. 143, https://web.stanford.edu/~hastie/StatLearnSparsity_files/SLS.pdf. [2]

Koenker, R. (2017), “Quantile regression: 40 years on”, *Annual Review of Economics*, Vol. 9/1, pp. 155-176, http://dx.doi.org/10.1146/annurev-economics-063016-103651. [10]

Koenker, R. (2005), *Quantile Regression*, Cambridge University Press, Cambridge, http://dx.doi.org/10.1017/CBO9780511754098. [11]

Koenker, R. and G. Bassett (1978), “Regression quantiles”, *Econometrica: The Journal of the Econometric Society*, Vol. 46/1, pp. 33-50, http://dx.doi.org/10.2307/1913643. [12]

OECD (2009), *PISA Data Analysis Manual: SAS, Second Edition*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264056251-en. [14]

StataCorp (2019), *Stata Statistical Software: Release 16.1*, StataCorp LLC, College Station, TX. [9]

Tibshirani, R. (1996), “Regression shrinkage and selection via the lasso”, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58/1, pp. 267-288, http://dx.doi.org/10.1111/j.2517-6161.1996.tb02080.x. [3]

Wang, H. et al. (2019), “Precision lasso: Accounting for correlations and linear dependencies in high-dimensional genomic data”, *Bioinformatics*, Vol. 35/7, pp. 1181-1187, http://dx.doi.org/10.1093/bioinformatics/bty750. [4]

Zhao, P. and B. Yu (2006), “On model selection consistency of lasso”, *Journal of Machine Learning Research*, Vol. 7/90, pp. 2541-2563, https://www.jmlr.org/papers/volume7/zhao06a/zhao06a.pdf. [5]

Notes

[1] TALIS-PISA link: Teaching and Learning International Survey (TALIS) and Programme for International Student Assessment (PISA) link covers schools that participated in both TALIS and PISA.

[2] For linear models, the goodness of fit measures that can be used to assess the performance of an estimator include the mean squared error (MSE), the R^2 and the Bayes information criterion (BIC), while, for non-linear models, such as logit, probit, and poisson models, such measures include the deviance and the deviance ratio.

[3] Apart from the 85% (training sample) versus 15% (validation sample) split, lasso regressions were estimated using 75% versus 25%, 95% versus 5% and even a no-split scenario with all observations used for the training sample. The sensitivity analysis showed that covariates selected by lasso were robust to the different sample splits. Indeed, the lists of covariates that got selected for each sample split were almost identical.

[4] Depending on how teacher variables are aggregated for a given model, there are four different balanced samples used for the analyses within this report – one each for: teacher variables averaged for all teachers, teacher variables averaged only for reading teachers, teacher variables averaged only for mathematics teachers and teacher variables averaged only for science teachers.

[5] In the case of quantile regressions conducted in Chapter 4, these controls are excluded due to the computational process (i.e. non-convergence of the quantile regression model). Indeed, the estimates for the models including controls for classmates' characteristics were missing for certain countries. Therefore, the models that include controls for classmates' characteristics are neither reported nor commented on. In the case of the analysis of within-school gender gaps in student performance, controls for classmates' characteristics are excluded due to the inclusion of school-level average student characteristics.

[6] Among the various different types of pseudo-R^2, this report applies McFadden's pseudo-R^2.

[7] Generating plausible values on an education test consists of drawing random numbers from the posterior distributions. For more detail on plausible values in general and on how to perform analyses with plausible values, see the *PISA Data Analysis Manual* (OECD, 2009[14]).

[8] Since Viet Nam does not have data on PISA test scores, it is not included in the analyses presented in Chapters 2 and 4.

[9] In Malta, there are only 17 out of the 44 schools that are not single-gender schools (i.e. all students surveyed in the school are same-gender students) and where the within-school differences in performance between girls and boys can be computed. Thus, it is not included in the school-level analysis presented in Chapter 4.

Annex C. List of tables available on line

The following tables are available in electronic form only.

Chapter 2. What do teachers and schools do that matters most for student achievement?

Table 2.51	Relationship between mathematics performance and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 2.52	Relationship between mathematics performance and classroom characteristics (focusing only on mathematics teachers and controlling for student characteristics)
Table 2.53	Relationship between science performance and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 2.54	Relationship between science performance and classroom characteristics (focusing only on science teachers and controlling for student characteristics)
Table 2.55	Relationship between reading performance and school culture (considering all teachers and controlling for student characteristics)
Table 2.56	Relationship between reading performance and school culture (focusing only on reading teachers and controlling for student characteristics)
Table 2.57	Relationship between reading performance and school culture (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.58	Relationship between reading performance and school culture (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 2.59	Relationship between mathematics performance and school culture (considering all teachers and controlling for student characteristics)
Table 2.60	Relationship between mathematics performance and school culture (focusing only on mathematics teachers and controlling for student characteristics)
Table 2.61	Relationship between mathematics performance and school culture (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.62	Relationship between mathematics performance and school culture (focusing only on mathematics teachers and controlling for student characteristics and classroom composition)
Table 2.63	Relationship between science performance and school culture (considering all teachers and controlling for student characteristics)
Table 2.64	Relationship between science performance and school culture (focusing only on science teachers and controlling for student characteristics)
Table 2.65	Relationship between science performance and school culture (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.66	Relationship between science performance and school culture (focusing only on science teachers and controlling for student characteristics and classroom composition)
Table 2.67	Relationship between science performance and induction (considering all teachers and controlling for student characteristics)
Table 2.68	Relationship between science performance and induction (focusing only on science teachers and controlling for student characteristics)
Table 2.69	Relationship between science performance and induction (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.70	Relationship between science performance and induction (focusing only on science teachers and controlling for student characteristics and classroom composition)
Table 2.71	Relationship between reading performance and type of professional development (considering all teachers and controlling for student characteristics)
Table 2.72	Relationship between reading performance and type of professional development (focusing only on reading teachers and controlling for student characteristics)
Table 2.73	Relationship between reading performance and type of professional development (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.74	Relationship between reading performance and type of professional development (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 2.75	Relationship between mathematics performance and type of professional development (considering all teachers and controlling for student characteristics)
Table 2.76	Relationship between mathematics performance and type of professional development (focusing only on mathematics teachers and controlling for student characteristics)
Table 2.77	Relationship between mathematics performance and type of professional development (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.78	Relationship between mathematics performance and type of professional development (focusing only on mathematics teachers and controlling for student characteristics and classroom composition)
Table 2.79	Relationship between reading performance and content of professional development (considering all teachers and controlling for student characteristics)

Table 2.80	Relationship between reading performance and content of professional development (focusing only on reading teachers and controlling for student characteristics)
Table 2.81	Relationship between reading performance and content of professional development (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.82	Relationship between reading performance and content of professional development (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 2.83	Relationship between reading performance and feedback received (considering all teachers and controlling for student characteristics)
Table 2.84	Relationship between reading performance and feedback received (focusing only on reading teachers and controlling for student characteristics)
Table 2.85	Relationship between reading performance and feedback received (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.86	Relationship between reading performance and feedback received (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 2.87	Relationship between mathematics performance and employment status (considering all teachers and controlling for student characteristics)
Table 2.88	Relationship between mathematics performance and employment status (focusing only on mathematics teachers and controlling for student characteristics)
Table 2.89	Relationship between mathematics performance and employment status (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.90	Relationship between mathematics performance and employment status (focusing only on mathematics teachers and controlling for student characteristics and classroom composition)
Table 2.91	Relationship between reading performance and school leadership (considering all teachers and controlling for student characteristics)
Table 2.92	Relationship between reading performance and school leadership (focusing only on reading teachers and controlling for student characteristics)
Table 2.93	Relationship between reading performance and school leadership (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.94	Relationship between reading performance and school leadership (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 2.95	Relationship between mathematics performance and school leadership (considering all teachers and controlling for student characteristics)
Table 2.96	Relationship between mathematics performance and school leadership (focusing only on mathematics teachers and controlling for student characteristics)
Table 2.97	Relationship between mathematics performance and school leadership (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.98	Relationship between mathematics performance and school leadership (focusing only on mathematics teachers and controlling for student characteristics and classroom composition)
Table 2.99	Relationship between science performance and school leadership (considering all teachers and controlling for student characteristics)
Table 2.100	Relationship between science performance and school leadership (focusing only on science teachers and controlling for student characteristics)
Table 2.101	Relationship between science performance and school leadership (considering all teachers and controlling for student characteristics and classroom composition)
Table 2.102	Relationship between science performance and school leadership (focusing only on science teachers and controlling for student characteristics and classroom composition)

Chapter 3. What do teachers and schools do that matters most for students' social and emotional development?

https://doi.org/10.1787/888934224258

Table 3.1	Between-school variance in students' social and emotional development, by non-cognitive outcome (considering all teachers)
Table 3.2	Variance in student perception of classroom disciplinary climate explained at the school level by teacher and school factors (considering all teachers)
Table 3.3	Variance in student perception of classroom disciplinary climate explained at the school level by teacher and school factors (focusing only on reading teachers)

Table 3.63	Relationship between student perception of teacher enthusiasm and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 3.64	Relationship between student perception of teacher enthusiasm and teacher well-being and job satisfaction (focusing only on reading teachers and controlling for student characteristics)
Table 3.65	Relationship between student perception of teacher enthusiasm and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics and classroom composition)
Table 3.66	Relationship between student perception of teacher enthusiasm and teacher well-being and job satisfaction (focusing only on reading teachers and controlling for student characteristics and classroom composition)
Table 3.67	Relationship between student perception of difficulty of the PISA test and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 3.68	Relationship between student perception of difficulty of the PISA test and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics and classroom composition)
Table 3.69	Relationship between student educational expectations and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 3.70	Relationship between student educational expectations and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics and classroom composition)

Chapter 4. Does what teachers and schools do matter differently depending on student achievement and gender?

https://doi.org/10.1787/888934224277

Table 4.1	Relationship between low-achieving students' performance in reading and classroom practices (considering all teachers and controlling for student characteristics)
Table 4.2	Relationship between high-achieving students' performance in reading and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.3	Relationship between low-achieving students' performance in mathematics and classroom practices (considering all teachers and controlling for student characteristics)
Table 4.4	Relationship between high-achieving students' performance in mathematics and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.5	Relationship between low-achieving students' performance in science and classroom practices (considering all teachers and controlling for student characteristics)
Table 4.6	Relationship between high-achieving students' performance in science and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.7	Relationship between low-achieving students' performance in reading and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 4.8	Relationship between high-achieving students' performance in reading and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.9	Relationship between low-achieving students' performance in mathematics and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 4.10	Relationship between high-achieving students' performance in mathematics and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.11	Relationship between low-achieving students' performance in science and teacher well-being and job satisfaction (considering all teachers and controlling for student characteristics)
Table 4.12	Relationship between high-achieving students' performance in science and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.13	Relationship between low-achieving students' performance in reading and working hours (considering all teachers and controlling for student characteristics)
Table 4.14	Relationship between high-achieving students' performance in reading and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.15	Relationship between low-achieving students' performance in mathematics and working hours (considering all teachers and controlling for student characteristics)
Table 4.16	Relationship between high-achieving students' performance in mathematics and classroom characteristics (considering all teachers and controlling for student characteristics)
Table 4.17	Relationship between low-achieving students' performance in science and working hours (considering all teachers and controlling for student characteristics)
Table 4.18	Relationship between high-achieving students' performance in science and classroom characteristics (considering all teachers and controlling for student characteristics)

Annex B. Technical notes on analyses in this report

https://doi.org/10.1787/888934224296

Table A B.1	Teacher and school factors included in the analyses
Table A B.2	Unweighted counts of schools, by number of sampled subject domain teachers
Table A B.3	Unweighted counts of students who repeated a grade, by sampled students' ISCED level
Table A B.4	Unweighted counts of schools, by sampled students' ISCED level