OECD Education Working Papers No. 188

The science of teaching science: An exploration of science teaching practices in PISA 2015

Tarek Mostafa,

Alfonso Echazarra,

Hélène Guillou

https://dx.doi.org/10.1787/f5bd9e57-en

OECD

Organisation for Economic Co-operation and Development

EDU/WKP(2018)24

Unclassified **English text only**

The science of teaching science: An exploration of science teaching practices in PISA 2015

OECD Education Working Paper No. 188

By Tarek Mostafa, Alfonso Echazarra and Hélène Guillou

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

Tarek Mostafa, Analyst, Early Childhood and Schools Division (tarek.mostafa@oecd.org).
Alfonso Echazarra, Analyst, Early Childhood and Schools Division (alfonso.echazarra@oecd.org).
Hélène Guillou, Research Assistant, Early Childhood and Schools Division (helene.guillou@oecd.org)

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

OECD EDUCATION WORKING PAPERS SERIES

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the authors.

Working Papers describe preliminary results or research in progress by the authors and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcome, and may be sent to the Directorate for Education and Skills, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

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

Lithuania was not an OECD member at the time of preparation of this publication. Accordingly, Lithuania is not included in the OECD average.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

---

www.oecd.org/edu/workingpapers

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Copyright © OECD 2018

ACKNOWLEDGEMENTS

The authors would like to thank Andreas Schleicher, Yuri Belfali, Miyako Ikeda, Francesco Avvisati, Pablo Fraser, Noémie Le Donné, Jeffrey Mo and the country representatives for valuable feedback on earlier drafts of this paper.

This paper was prepared by Tarek Mostafa, Alfonso Echazarra and Hélène Guillou from the OECD Directorate for Education and Skills, and edited by Marilyn Achiron.

ABSTRACT

This paper explores the relationship between various science teaching strategies and students' science-related outcomes. The focus is on enquiry-based science teaching, teacher-directed instruction, adaptive teaching and teacher feedback. The outcomes of interest include students' science performance, and students' dispositions and attitudes towards science.

The findings show that the negative association between enquiry-based science teaching and science performance is greatly attenuated when lessons are delivered in disciplined science classes. This approach could help close the gender gap between girls and boys when it comes to attitudes towards science and to the decision to pursue a career in STEM-related fields. The results also show that teacher-directed instruction is a reliable strategy that is positively associated with students' science outcomes regardless of school climate and resources. Adaptive teaching is positively correlated with science performance in the majority of countries, particularly in countries known for the use of personalised learning approaches, while teacher feedback is weakly but positively associated with science performance once students' achievement in mathematics and reading is accounted for. In general, all teaching strategies have the potential to foster enjoyment of and interest in science, and students' epistemic beliefs, self-efficacy in science and expectations of a career in science.

RÉSUMÉ

Cet article explore la relation entre diverses stratégies d'enseignement des sciences et les résultats des étudiants dans cette matière. L'accent est mis sur l'enseignement fondé sur l'investigation, l'enseignement dirigé par l'enseignant, l'enseignement différencié et le feedback des enseignants. L'article s'intéresse aux performances des élèves en science, ainsi que leurs dispositions et leurs attitudes à l'égard des sciences.

Les résultats montrent que l'association négative entre l'enseignement des sciences basé sur l'investigation et la performance en sciences est largement atténuée lorsque les leçons se déroulent dans un environnement discipliné. Cette approche pourrait aussi aider à réduire l'écart entre les filles et les garçons en ce qui concerne les attitudes à l'égard des sciences et la décision de poursuivre une carrière dans les domaines liés aux STIM (sciences, technologie, ingénierie et mathématiques). Les résultats montrent également que l'enseignement dirigé par l'enseignant est une stratégie fiable, associée positivement aux résultats scientifiques des élèves, indépendamment de l'environnement et des ressources de l'école. L'enseignement différencié est positivement corrélé aux performances scientifiques dans la majorité des pays, en particulier dans les pays connus pour l'utilisation d'approches d'apprentissage personnalisées, tandis que le feedback des enseignants est associé de manière faible mais positive aux performances scientifiques, une fois que les résultats des élèves en mathématiques et en compréhension de l'écrit ont été pris en compte. En général, toutes les stratégies d'enseignement ont le potentiel de favoriser l'intérêt des élèves pour les sujets scientifiques, leurs convictions épistémiques, leur efficacité perçue en sciences et leurs aspirations à embrasser une carrière scientifique.

Table of contents

ACKNOWLEDGEMENTS **3**

ABSTRACT **4**

1. Introduction **8**

2. PISA 2015 data and the methodological approach **11**

3. Enquiry-based science teaching **14**

3.1. Evidence of the use of enquiry-based science teaching 18
3.2. Enquiry-based science teaching and students' attitudes and expectations 22
3.3. Enquiry-based science teaching and performance in science 28

4. Teacher-directed science instruction **40**

4.1. Evidence of the use of teacher-directed science instruction 41
4.2. Teacher-directed science instruction and students' attitudes and expectations 46
4.3. Teacher-directed science instruction and performance in science 51

5. Adaptive instruction in science lessons **57**

5.1. Descriptive evidence on the use of adaptive teaching in science lessons 58
5.2. Adaptive teaching in science lessons and students' attitudes and expectations 61
5.3. Adaptive teaching in science lessons and performance in science 64
5.4. What motivates science teachers to adapt their teaching? 65

6. Feedback in science classes **74**

6.1. Descriptive evidence on the use of teacher feedback in science lessons 76
6.2. Teacher feedback in science lessons and students' attitudes and expectations 80
6.3. Teacher feedback in science lessons and performance in science 83

7. What the results imply for policy and practice **90**

Choose teaching strategies suitable for the school context. 90
Foster discipline to ensure the success of enquiry-based science teaching. 90
Teacher-directed science instruction is always a reliable strategy regardless of the school context. 91
Use adaptive teaching and teacher support to enhance the learning experience of diverse student populations. 91
Teacher feedback complements other strategies and could help the low performers. 92
All teaching practices are effective in improving students' attitudes and dispositions towards science. 92

References **93**

Notes **102**

Annex A **103**

Science test items 103
Science subscales 103
Quality assurance 105
Missing data 105
Quartiles 106

Odds ratios ... 106
Standard errors and significance tests ... 106
Causation vs. correlation ... 107

Annex B ... 108

Student-level indices ... 108
School-level indices ... 111

Annex C ... 112

List of tables ... 112

Figures

Figure 3.1. Index of enquiry-based science teaching, reported by students ... 20
Figure 3.2. Index of enquiry-based science teaching, reported by teachers ... 21
Figure 3.3. Enquiry-based teaching in science lessons ... 22
Figure 3.4. Enjoyment of science and enquiry-based science teaching ... 23
Figure 3.5. Science self-efficacy and enquiry-based science teaching ... 24
Figure 3.6. Epistemic beliefs in science and enquiry-based science teaching ... 26
Figure 3.7. Expectations of a science-related career and enquiry-based science teaching ... 27
Figure 3.8. Performance in science and enquiry-based science teaching ... 29
Figure 3.9. Performance in science and enquiry-based science teaching, by disciplinary climate ... 31
Figure 3.10. Science performance and index of enquiry-based science teaching, by performance in science ... 34
Figure 3.11. Enquiry-based science teaching and success in science items ... 35
Figure 3.12. Teaching practices and expectations of pursuing a science-related career ... 36
Figure 3.13. Teaching practices and science-related attitudes ... 37
Figure 3.14. Enquiry-based teaching practices and expectations of a science-related career ... 38
Figure 3.15. Enquiry-based science teaching and expectation of working in different science-related careers ... 39
Figure 4.1. Index of teacher-directed science instruction, reported by students ... 43
Figure 4.2. Index of teacher-directed instruction, reported by teachers ... 44
Figure 4.3. Teacher-directed instruction in science lessons ... 45
Figure 4.4. Relationship between teacher-directed science instruction and enquiry-based science teaching ... 46
Figure 4.5. Enjoyment of science, interest in science and teacher-directed science instruction ... 48
Figure 4.6. Epistemic beliefs, science self-efficacy and teacher-directed science instruction ... 49
Figure 4.7. Expectations of a science-related career and teacher-directed science instruction ... 51
Figure 4.8. Performance in science and teacher-directed science instruction ... 52
Figure 4.9. Science teaching practices and student performance in science ... 53
Figure 4.10. Teacher-directed science instruction and student success on science items ... 55
Figure 5.1. Index of adaptive teaching in science lessons, reported by students ... 59
Figure 5.2. Adaptive teaching in science lessons ... 61
Figure 5.3. Enjoyment of science and adaptive teaching in science lessons ... 62
Figure 5.4. Epistemic beliefs in science and adaptive teaching in science lessons ... 63
Figure 5.5. Expectations of a science related-career and adaptive teaching in science lessons ... 64
Figure 5.6. Performance in science and adaptive teaching in science lessons ... 65
Figure 5.7. Stratification and adaptive teaching in science lessons ... 66
Figure 5.8. Association between science performance and adaptive teaching by education systems' stratification ... 67

Figure 5.9. School size and adaptive teaching in science lessons 68
Figure 5.10. Disciplinary climate and adaptive teaching in science lessons 69
Figure 5.11. Student diversity at school and adaptive teaching in science lessons 70
Figure 5.12. Performance in science and teacher support in science lessons 73
Figure 6.1. Index of teacher feedback in science lessons, reported by students 77
Figure 6.2. Gender differences in the frequency of science teachers' feedback 79
Figure 6.3. Teacher feedback in science lessons 80
Figure 6.4. Enjoyment of science and teacher feedback in science lessons 81
Figure 6.5. Epistemic beliefs in science and teacher feedback in science lessons 82
Figure 6.6. Expectations of a science-related career and teacher feedback in science lessons 83
Figure 6.7. Performance in science and teacher feedback in science lessons 84
Figure 6.8. Teacher feedback practices and science performance 85
Figure 6.9. Telling students how they are performing in the science course and science performance 86
Figure 6.10. Index of teacher feedback and adjusted science performance 87
Figure 6.11. Teaching practices by class size 88
Figure 6.12. Teaching practices by educational level and programme orientation 89

Boxes

Box 2.1. Construction of indices on science teaching practices based on teacher-reported data 12
Box 3.1. Enquiry-based science teaching 17
Box 3.2. Teacher-reported use of EBST and student outcomes 30
Box 3.3. Using technology to support enquiry-based science teaching 32
Box 3.4. Getting boys and girls engaged in science 36
Box 4.1. Teacher-directed science instruction 41
Box 4.2. Teacher-reported use of TDSI and student outcomes 50
Box 5.1. Adaptive teaching in science lessons 58
Box 5.2. Teacher support 71
Box 6.1. Teacher feedback in science lessons 76
Box 6.2. Do smaller classes allow for certain teaching practices more than larger classes? 87

1. Introduction

Science permeates all aspects of modern life. It is all around us, from the humble toaster to the mighty rocket putting satellites into orbit. Science's record in improving our living circumstances through medicine, communication, transport and many other fields is undeniable.

In today's world, proficiency in science is not a luxury but a necessity. According to the United States Bureau of Labour Statistics, in 2015, 8.6 million jobs in the United States (representing 6.2% of all jobs) were in science-, technology-, engineering- and mathematics-related fields (Fayer, Lacey and Watson, 2017). Jobs in science and mathematics, in particular, are expected to grow at an unprecedented rate of 28.2% between 2014 and 2024, compared to 6.5% growth in all other professions. This rise will be accompanied by the progressive automation of routine and low-skilled jobs. Figures from the World Bank show that a wide range of jobs – from truck drivers to finance professionals – have a high probability of being automated in the coming years, with technology entirely or largely replacing routine tasks performed by human workers. This evidence underscores the importance of science in the future. Students who perform well in science are more likely to pursue careers in this field and more likely to find good jobs (OECD, 2016a). Similarly, at the macro level, countries with more scientists will be more competitive and will enjoy greater growth and labour productivity (World Bank, 2016).

In previous research conducted with data from PISA 2015, performance and attitudes towards science were examined in light of student background and school characteristics (OECD, 2016b). The evidence highlighted the importance of the disciplinary climate in school, time spent studying, the absence of early selection of students into different programmes (tracking), and the availability and equitable allocation of resources, for fostering student achievement in science. However, little is known about the ways in which teaching practices are related to student science performance.

Recent studies have emphasised the importance of teachers for learning. However, the question of what makes a teacher successful in improving students' outcomes has not been settled yet (Aaronson, Barrow and Sander, 2007). Existing literature focuses on a range of teacher-related characteristics. For instance, some studies assessed the impact of teachers' race and gender (Dee, 2005, 2007), while others focused on teachers' qualifications (Kane, Rockoff and Staiger, 2008). However, these observable and easily measured variables are rarely found to be correlated with student achievement; when they are, they explain a modest fraction of variations in performance (Rivkin, Hanushek and Kain, 2005). This has led to a growing interest in what teachers actually do in the classroom as opposed to their background.

Several studies indicate that instructional practices could have a more significant effect on students' science performance and attitudes than teachers' experience and advanced degrees (Kloser, 2014; Rockoff, 2004; Seidel and Shavelson, 2007). Indeed, what teachers enact in the classroom has the potential to engage students with science or alienate them from it. This, in turn, highlights the need to identify the core teaching practices that have a positive impact on students' science performance and attitudes.

This paper examines the relationships between several teaching practices and students' cognitive and non-cognitive outcomes in science. Cognitive outcomes cover achievement in science and non-cognitive outcomes include students' attitudes, epistemic beliefs and

career ambitions. This paper aims to contribute to the growing body of literature focusing on the effect of teaching practices on student performance (Gallimore et al., 2009; Loewenberg Ball and Forzani, 2009; Windschitl, Thomspon and Braaten, 2008). In particular, it expands the analyses previously conducted on mathematics teaching practices using PISA 2012 data (Echazarra et al., 2016).

Before examining the relationship between science teaching practices and student outcomes it is important to define what a teaching practice is. (Reckwitz, 2002) defines a practice as a routinised behaviour that consists of a number of elements involving tools, motivation and physical and mental activities. (Spillane, 2012), on the other hand, defines a teaching practice as the co-ordinated, patterned and meaningful interaction between people.

In both definitions, a strategy involves complex and multifaceted routines and patterns of interaction (Kloser, 2014). Such routines encompass two forms. The first is the ostensive form, based on an idealised abstract understanding of what the practice consists of and what it should lead to; something that makes the practice identifiable in a classroom setting. The second is the performative form, describing what takes place in a real, complex classroom environment (Feldman and Pentland, 2003). The performative form varies from one classroom to another, and depends on the characteristics of students, teachers and schools.

The ostensive form of the practice is something that teachers encounter in a textbook about pedagogies or in a teacher training course, while the performative form is what they encounter in a real classroom. A successful practice is one that can bridge the gap between the two forms by taking into account the level of student engagement, their culture and characteristics, and the organisational structure of a school. In this sense, a practice can be thought of as a socio-cultural act (Vygotskij and Cole, 1981). It is not only what teachers do, but also how teachers interact with their environment, with different school actors, and how they take the diversity of students into account (Cook and Brown, 1999).

From a survey point of view, the distinction between the two forms is important. When students and teachers are asked to report on the teaching practices they experience or implement in a classroom, perceptions tend to vary. Students' experiences tend to converge towards the performative form of the practice because this is what takes place in a classroom, while teachers tend to be more aware of the ostensive nature of what they are implementing. This raises the need to survey both students and teachers, and to adapt the survey questions to a particular discipline, even though some core teaching practices might have a cross-disciplinary nature (Kloser, 2014). In PISA 2015, science teaching practices were assessed from the perspective of both the student and the teacher.

The findings highlight the role of classroom discipline in the success of enquiry-based science teaching. This teaching approach could also help close the gender gap in some countries by nurturing girls' science-related achievements and attitudes. Teacher-directed instruction is found to be a reliable teaching strategy whose effectiveness is not sensitive to the surrounding school environment. Moreover, adaptive teaching, teacher feedback and support are found to improve science performance although in some cases the effects are moderate. In general, all teaching strategies have the potential to improve student interest in and enjoyment of science, epistemic beliefs, self-efficacy and expectations of a career in science.

The paper is organised as follows. Section 2 presents the PISA 2015 data, the teacher survey, and the methodologies used in the paper. Section 3 explores the relationship between enquiry-based science teaching and student outcomes. Section 4 focuses on

teacher-directed science instruction. Section 5 examines adaptive teaching. Section 6 investigates teacher support and feedback in science classes. The concluding section presents the implications of the findings. All supplementary materials are found in the annexes.

2. PISA 2015 data and the methodological approach

In 2015 the Programme for International Student Assessment (PISA) collected data on more than 540 000 students representing about 29 million 15-year-olds in 72 participating countries and economies. The key area of assessment was science; reading, mathematics and collaborative problem solving were minor areas of assessment.

Science assessments were carried out using a mixture of multiple-choice questions and questions requiring students to construct their own responses. All tests were computer based and lasted a total of two hours for each student. [1] About 810 minutes of test items for science, reading, mathematics and collaborative problem solving were covered, with different students taking different combinations of test items.

A background questionnaire was filled out by students and took approximately 35 minutes to complete. The questionnaire sought information about the students, their home, and their school and learning experiences. In particular, the student questionnaire collected data on students' perception of teaching practices used in science classes. The surveyed teaching practices covered enquiry-based science teaching, teacher-directed instruction, teacher support, teacher feedback and adaptive teaching.

In addition to the student background questionnaire, PISA 2015 – for the first time in the history of the assessment – distributed an optional questionnaire to teachers. Nineteen participating countries and economies (nine OECD countries and ten partner countries and economies) opted to distribute the questionnaire. The questionnaire was intended to provide contextual information on the teachers of typical 15-year-old students eligible to participate in the PISA study. In particular, teachers provided information on the science teaching practices they use, such as enquiry-based science teaching, teacher-directed instruction, and team work in science classes. [2]

Teachers were sampled as part of two populations: science and non-science teachers. Since the focus of this paper is on science teaching strategies, all analyses with teacher data in this paper were carried out with the sample of science teachers since they were the ones answering science-related questions. Moreover, students and teachers in PISA 2015 were sampled randomly and independently within each school. In other words, it is not possible to determine whether an individual teacher is teaching a particular student. In order to analyse student and teacher data jointly, teacher-reported data had to be aggregated at the school level. Therefore, any teacher-level variable should be interpreted as a school average of what the science teachers within each school reported. For a detailed description of the sampling procedures, the content of the teacher questionnaire, and the aggregation procedure, see (OECD, 2017a), and (Mostafa and Pál, 2018).

Box 2.1. Construction of indices on science teaching practices based on teacher-reported data

This paper relies on a number of composite indices developed using student- and teacher-reported data on science teaching practices. Using student-reported data, five indices were constructed and were made available in the published version of the PISA 2015 dataset. However, no indices were constructed using the teacher-reported data.

Therefore, based on teacher data two new indices (question TC037) were created. The indices are: enquiry-based science teaching and teacher-directed instruction.

The three indices were constructed as the sum of the items constituting each index (the items were coded on a four-point scale ranging from 1 to 4). The sum was averaged at the school level and merged with the student data before being standardised to have an average of 0 and a standard deviation of 1 for the nine OECD countries. In this sense, the teacher-reported science teaching practices should be understood as school attributes representing an aggregate of what the teachers within each school reported.

This paper relies on a variety of methodologies to explore the relationship between teaching strategies and science outcomes. In particular, four types of statistics are presented: descriptive statistics, and simple, logistic, fixed effects, and quantile regressions.

Descriptive statistics, such as means and proportions, are used to describe the extent to which students and teachers experience and implement particular teaching practices. Cross-tabulations of these practices by student and school characteristics are used to provide a description of the context in which the practices are used.

Simple regressions (OLS) are used to measure the strength of the association between teaching practices and continuous science-related outcomes, including PISA scores and values on the indices, such as enjoyment and interest in science.

Fixed-effects models are used to account for student and teacher selection into schools. For instance, teachers with certain abilities might select or be selected for schools with an inclination for particular pedagogical practices (e.g. Montessori schools). This results in biased findings if unobserved school characteristics are not accounted for. Fixed-effects models allow the researcher to correct for such bias. These models should be understood as a generalised form of simple regressions, where all school characteristics (observed and unobserved) are accounted for.

Logistic regressions are used to analyse the difficulty of test items. PISA 2015 used 184 science test items to measure students' performance. These items were of varying difficulty, designed to assess the performance of students with different levels of ability. High-performing students were expected to answer most questions correctly; low-performing students would only answer the easy items correctly. The answer to each test item was coded as correct or incorrect (i.e. binary). A logistic model is used to explore how teaching practices affect the likelihood of a student giving a correct answer to each item. This analysis shows whether a particular teaching practice is better suited for helping students engage with more difficult science tasks and concepts. Logistic regressions are also used for analysing science-related career choices. A binary outcome – whether or not

students expect to pursue a science-related career – is related to teaching practices. Two versions of these regressions are estimated: univariate and multivariate, as discussed below.

Quantile regressions are used to examine differences in the relationship between teaching practices and science scores among students with different levels of performance. In other words, it is possible for a teaching practice to have stronger associations with performance among top performers than among low performers. Such differences can be identified using this method.

To facilitate the interpretation of regression coefficients, a convention is adopted where coefficients are classified as weak, moderate or strong, depending on their magnitude. The cut-off points for the three categories are 20- and 30-point changes in science scores, and 0.2 and 0.3 of a unit change on scale indices, where less than 20 points (or 0.2 of a unit change for scale indices) is weak, 20 to 30 points (or 0.2 to 0.3 of a unit change for scale indices) is moderate, and higher than 30 points (or 0.3 of a unit change for scale indices) is strong.

Most regression analyses are carried out using two specifications. The first is univariate, with only one variable included as an independent variable (i.e. teaching practice). The second is multivariate, with teaching practice included among a number of student and school characteristics (i.e. student socio-economic status, gender, grade, number of science courses attended, and average school socio-economic profile in addition to the teaching strategy of interest).

All scale indices used in this paper (e.g. economic, social and cultural status [ESCS]) were constructed using answers to PISA questionnaire items as described in Annex B.

3. Enquiry-based science teaching

Around the world, substantial resources have been invested in improving science education, in reforming curricula, and in building science teachers' skills. In particular, teachers have been encouraged to use enquiry in their instruction of science content. In the United States, the National Science Foundation, the American Association for the Advancement of Science and the National Research Council invested considerable resources towards the achievement of that goal. The (National Research Council (United States), 1996) education reform document advocated that science teachers should engage students in thinking about science as enquiry. This document described a range of instructional approaches, from open enquiry where students take the lead in identifying science problems, raise questions, design experiments, record observations and develop a solution to the problem, to more structured enquiry, where teachers define the topic and procedures to follow.

Science enquiry first appeared in the debate over the nature of learning and teaching in the work of leading theorists like Jean Piaget, Lev Vygotsky and David Ausubel. Their work on the philosophy of learning was later known as constructivism (Cakir, 2008; Minner, Levy and Century, 2010).

Through scientific enquiry, students should develop a critical way of engaging with science. They should be able to acquire a deep understanding about a topic, develop a coherent scientific method, and ultimately provide a robust answer to the question under investigation (Crawford, 2007). However, the implementation of enquiry-based science teaching is fraught with challenges.

The first challenge relates to the definition of enquiry-based teaching. For instance, minimally guided discovery, project-based learning, and enquiry learning are sometimes lumped under the same heading even though the level of teacher involvement might differ from one practice to another. This results in the application of blanket criticism of strategies that, in practice, are very different from one another (Hmelo-Silver, Duncan and Chinn, 2007). The absence of a common definition, and the continuous evolution of that definition, highlight the challenges of determining what constitutes scientific enquiry (Duschl et al., 2007; Furtak et al., 2012).

When it comes to unguided discovery, criticism has focused on the lack of structure in the construction of knowledge. According to critics, novice learners do not have the extensive knowledge or training of professional scientists. When scientists formulate a hypothesis they draw on a body of knowledge built over a long period of time. In contrast, students lack this knowledge, and can only rely on a patchy understanding of scientific principals and on a short-term memory that could become overloaded with new information (J. Sweller, 2003, 2004). The increased load of information, in turn, prevents the accumulation of real knowledge (J. Sweller, van Merrienboer and Paas, 1998; John Sweller, 1999).

Another critique advanced against constructivist practices has focused on the shift of emphasis from learning the content of a discipline towards experiencing the procedures of that discipline (Handelsman et al., 2004) – a shift that could ultimately lead to a rejection of instruction based on the facts in favour of extensive practical work. According to (Kirschner, 1992; Kirschner, Sweller and Clark, 2006), this excessive focus on process neglects the differences between how science is practiced (epistemology) and how science is learned (pedagogy).

Moreover, learning science happens in a school context. Therefore, it is expected that the success of a certain teaching practice will depend on the contextual sensitivity of that practice. For instance, successful enquiry-based learning requires a positive school environment, discipline, equipment and personnel, sufficient instruction time, and a school leadership that encourages scientific enquiry in addition to well-trained teachers who are capable and willing to implement this strategy. In contrast, teacher-directed instruction might require less of these resources. The contextual sensitivity of enquiry-based learning also implies that what takes place in a classroom (the performative form of a practice) might diverge substantially from the abstract, or ostensive, form of the practice.

In addition, the implementation of enquiry-based science instruction requires the teacher to relinquish some control over the classroom in favour of the students. The success of this approach needs a different set of skills and attitudes than a teacher-driven lecture. In fact, a lecture is more akin to a scripted performance; enquiry-based instruction is more about improvisation and adaptation. In this sense, the successful adoption of these practices depends on teachers' capacity and willingness to enact enquiry-based teaching (McGinnis, Parker and Graeber, 2004; Newman et al., 2004), on teachers' attitudes towards the practices(Windschitl, 2003), and on the existence of a school culture, supported by parents, students and education authorities, that encourages scientific enquiry (McGinnis, Parker and Graeber, 2004).

The fact that enquiry-based teaching has not been largely adopted by science teachers could also be attributed to the lack of conclusive evidence on its positive effect on student outcomes. For instance, (Strijbos, Kirschner and Martens, 2004) and (Mayer, 2004) cast doubt on the efficacy of unguided enquiry-based teaching by arguing that such practices limit the role of the teacher and allow students to engage in self-guided activities of dubious value. (Mayer, 2004) reviewed a number of foundational experimental studies dating back to the 1960s and contrasted unguided enquiry approaches with teacher-guided instruction. Invariably, the author found that a guided approach was better at building knowledge without completely refuting the merits of enquiry when it is combined with teacher guidance. This evidence is corroborated in studies focusing on problem-based learning and enquiry learning (Hmelo-Silver, Duncan and Chinn, 2007). In both types of learning, students learn content and discipline-specific processes. Both approaches rely on authentic problems and questions and could be combined with teacher-directed instruction (Krajcik, Czerniak and Berger, 1999; Schwartz and Bransford, 1998). Positive evidence in favour of problem-based learning was found by (Dochy et al., 2003), and in favour of enquiry learning by a number of controlled experiments (Geier et al., 2008; Hickey, Wolfe and Kindfield, 2000; Lynch et al., 2005). A recent report by McKinsey also found evidence that students who receive a blend of enquiry-based and teacher-directed instruction have the best outcomes (Chen et al. 2017).

When it comes to PISA, the preliminary results of the 2015 round uncovered a negative correlation between enquiry-based science teaching and performance in science even after accounting for student and school socio-economic profiles (OECD, 2016b). On the other hand, using PISA 2012 data, (Echazarra et al., 2016) found that a combination of teacher-directed instruction, team work and cognitive activation were necessary to improve performance in mathematics.

In this section, PISA 2015 data is used to investigate the association between enquiry-based science teaching and student outcomes in science. Different methodologies are used in order to decipher the negative relationship described in Volume II of the PISA 2015 initial report (OECD, 2016b). In particular, the analyses consider both cognitive and non-

cognitive science outcomes instead of focusing only on the former. They explore the interactions between enquiry-based teaching and the school environment, differences in the benefits of such teaching practices, depending on student proficiency in science (e.g. top performers might benefit more from scientific enquiry) and on the branch of science being taught (e.g. chemistry, biology), and whether enquiry is more useful in helping students engage with more difficult science tasks.

Box 3.1. Enquiry-based science teaching

The use of enquiry-based teaching was measured in PISA using both student- and teacher-reported information.

For the students, one question (question ST098) with nine items was used. The question asked about the frequency with which certain enquiry-based practices are undertaken in science classes. Answers were provided on a four-point Likert scale ranging from "In all lessons", "In most lessons", "In some lessons", to "Never or hardly ever". The question is:

When learning <school science> topics at school, how often do the following activities occur?

1. Students are given opportunities to explain their ideas.
2. Students spend time in the laboratory doing practical experiments.
3. Students are required to argue about science questions.
4. Students are asked to draw conclusions from an experiment they have conducted.
5. The teacher explains how a <school science> idea can be applied to a number of different phenomena (e.g. the movement of objects, substances with similar properties).
6. Students are allowed to design their own experiments.
7. There is a class debate about investigations.
8. The teacher clearly explains the relevance of <broad science> concepts to our lives.
9. Students are asked to do an investigation to test ideas.

An index was constructed based on students' responses to these nine statements using IRT (Item Response Theory) scaling (see OECD, 2017c, Technical Report, Chapter 16). The index was standardised to have an average of 0 across OECD countries and a standard deviation of 1, meaning that two-thirds of the population fall between the values of -1 and 1 on the index.

For the teachers, one question (question TC037) with 22 items was used to measure the frequency of certain activities used in science classes. Eight items correspond to enquiry-based science practices. Answers were provided on a four-point Likert scale ranging from "Never or almost never", "Some lessons", "Many lessons", to "Every lesson or almost every lesson". The question is:

How often do these things happen in your <school science> lessons?

1. Students are asked to draw conclusions from an experiment they have conducted.
2. Students are given opportunities to explain their ideas.
3. Current scientific issues are discussed.
4. Students make calculations using scientific formulas.
5. Students do their own scientific study and related research.
6. Students carry out practical work.
7. Students write up laboratory reports.
8. I discuss questions of practical relevance.

These eight items were used to construct a teacher-reported index of use of enquiry-based teaching using the method described in Box 2.1. For both student- and teacher-reported indices, higher values on the index indicate greater reliance on enquiry-based science teaching practices.

The two indices based on student- and teacher-reported information rely on differently worded questionnaire items. They are also analysed at different levels. The student-reported index is measured at the student level while the teacher-reported index is analysed at the school level. In other words, the former contains both between-student and between-school variations while the latter only contains between-school variations. For these reasons, the two indices should not be compared.

3.1. Evidence of the use of enquiry-based science teaching

Figure 3.1 presents the average of the student-reported index of exposure to enquiry-based science teaching (hereafter "EBST"). The findings show that students in Albania, Algeria, Canada, Denmark, the Dominican Republic, Georgia, Indonesia, Jordan, Kosovo, Lebanon, Mexico, Moldova, Peru, Portugal, Qatar, the Russian Federation (hereafter "Russia"), Sweden, Tunisia, Turkey, the United Arab Emirates, and the United States experience EBST more frequently. All of these countries have an average on the index which is more than 0.25 of a standard deviation higher than the OECD average. In contrast, Austria, Beijing-Shanghai-Jiangsu-Guangdong (China) (hereafter "B-S-J-G [China]"), Finland, Japan, Korea, Netherlands, Spain and Chinese Taipei are among the countries with limited reliance on EBST.

Variations in student-reported experiences of EBST are the largest in (B-S-J-G) China, Bulgaria, Israel, Korea, Montenegro, Qatar, the Slovak Republic, Chinese Taipei, Turkey and the United Arab Emirates; they are the smallest in Albania, Ciudad Autónoma de Buenos Aires (Argentina) (hereafter "CABA [Argentina]"), Denmark, Indonesia, Ireland, Kosovo, Latvia, Lebanon, Macao (China), Moldova, Romania and Viet Nam.

When comparing student-reported experiences of EBST by school characteristics the findings show that in some countries EBST is more prevalent in schools with a socio-economically disadvantaged profile. In fact, in 27 countries EBST is more frequently experienced by students in disadvantaged schools (i.e. schools in the bottom 25% of the PISA index of economic, social and cultural status [ESCS]) compared with students in the top quarter of ESCS. The difference between top and bottom quarters is negative and statistically significant in 27 out of 70 participating countries and economies. The difference is particularly large in Bulgaria, Georgia, Greece, Israel, Italy, Montenegro, Peru, Qatar and the United Arab Emirates. The difference is non-significant in 32 countries and economies while the opposite is true in 11.

A possible explanation behind this finding is that EBST is more prevalent in vocational schools and tracks, which tend to be populated mostly by disadvantaged students. The exceptions to this result are Albania, (B-S-J-G) China, Denmark, Finland, France, Germany, Iceland, Japan, Malta, Singapore and Sweden, where EBST is more prevalent in advantaged schools (schools in the top 25% of ESCS).

A comparison between public and private schools in the prevalence of EBST reveals few significant differences. EBST is more prevalent in public schools in Canada, Japan, Korea, Peru, Poland, Qatar, Chinese Taipei, the United Arab Emirates and the United States; while

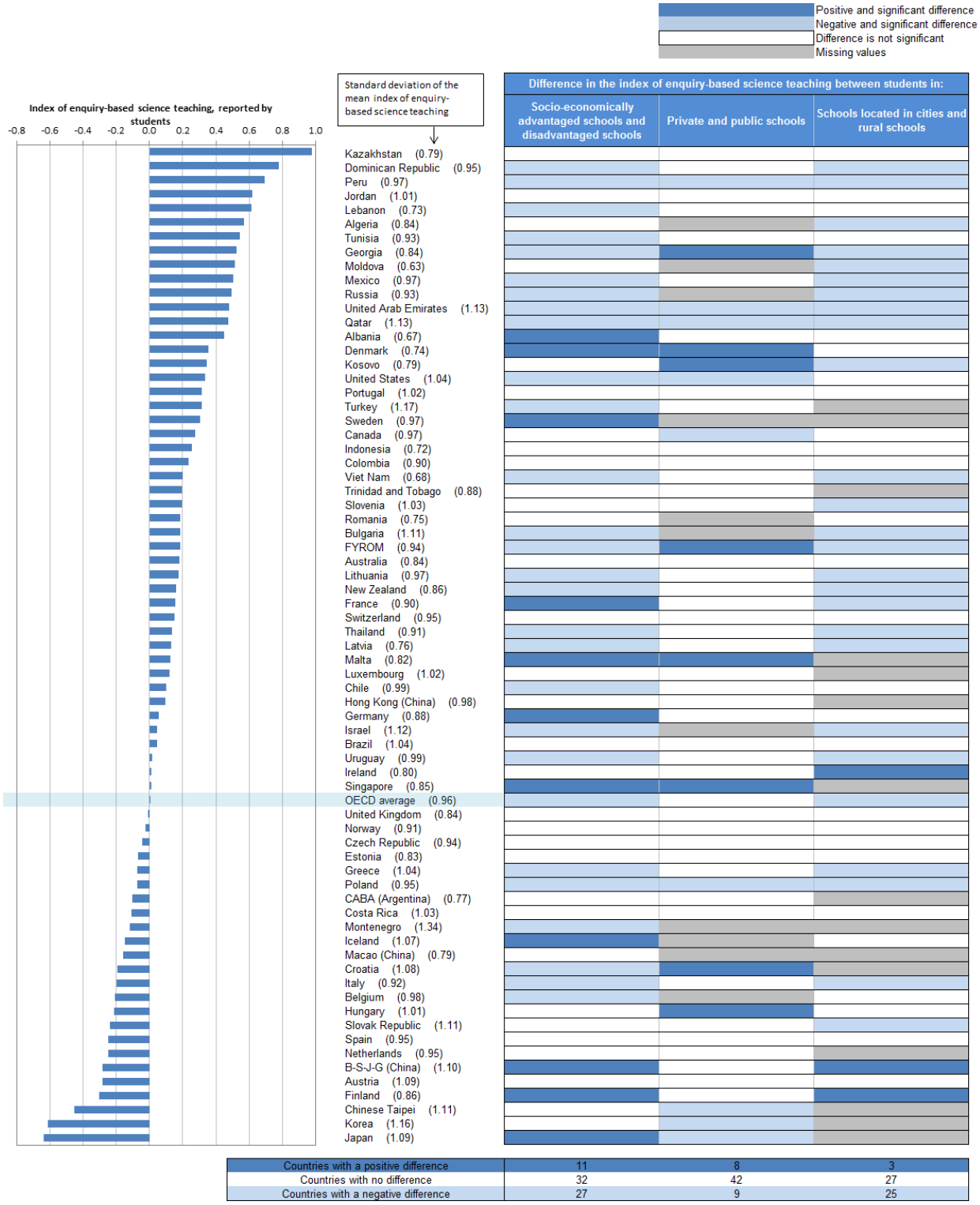
the reverse is true in Croatia, Denmark, the Former Yugoslav Republic of Macedonia (hereafter "FYROM"), Georgia, Hungary, Kosovo, Malta and Singapore.

When it comes to school location, the findings show that students in rural schools are more exposed to EBST than their peers living in cities with more than 100 000 inhabitants. This is the case in 25 participating countries and economies where differences are statistically significant. The differences between urban and rural schools are particularly large in Bulgaria, FYROM, Georgia, Greece, the Slovak Republic and Slovenia. Only students in rural schools in B-S-J-G China, Finland and Ireland reported less frequent exposure to EBST.

The concentration of immigrant students in a school appears to have little effect on students' exposure to EBST except in few countries (Table 3.2). In Canada, Greece, Luxembourg and Singapore, students in schools where more than 30% of their schoolmates have an immigrant background reported more frequent exposure to EBST; in Israel, Lithuania, New Zealand, Qatar and the United Arab Emirates, students in these types of schools reported less frequent exposure to EBST. Moreover, in Bulgaria, Finland, Italy, Lithuania, Macao (China), the Slovak Republic, Spain, Switzerland, Turkey and Uruguay, students in schools where more than 30% of their schoolmates speak a language different from the language of instruction reported more frequent exposure to EBST; only in Qatar and the United Arab Emirates did such students report less frequent exposure to EBST.

Based on these descriptive results, it is clear that student-reported experiences of EBST are related to student and school characteristics. In general, EBST is more prevalent in disadvantaged schools. From a methodological perspective, the effect of science teaching strategies on student outcomes is likely to be confounded by student and school profiles. Therefore, regression analyses should account for these confounders. In this paper, this is done by adopting a school fixed-effects approach and by accounting for students' socio-economic status.

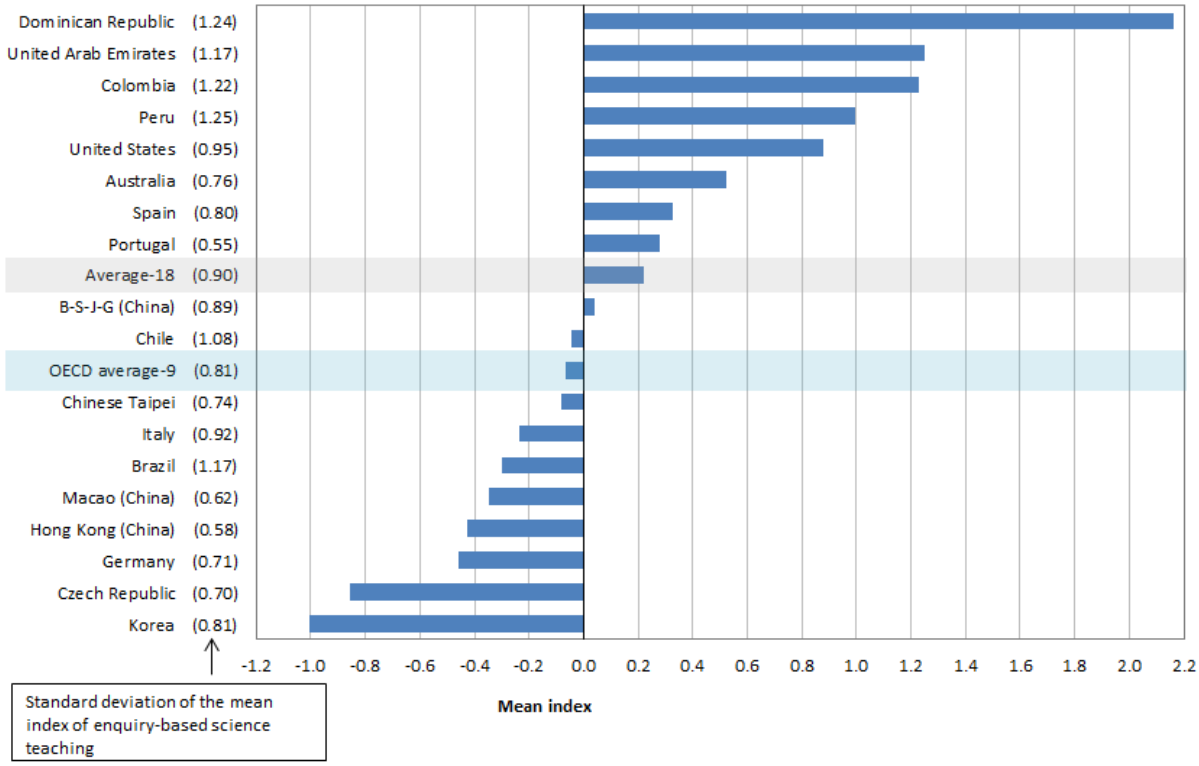
Figure 3.1. Index of enquiry-based science teaching, reported by students



Countries and economies are ranked in descending order of the index of enquiry-based science teaching, reported by students.
Source: OECD, PISA 2015 Database.

Based on teacher-reported data on the implementation of EBST, the findings show that these teaching practices are most commonly used, in descending order of frequency, in the Dominican Republic, the United Arab Emirates, Colombia, Peru, the United States, Australia, Spain and Portugal. The practices are the least frequently used in Korea, followed in descending order by the Czech Republic, Germany, Hong Kong (China), Macao (China), Brazil and Italy (Figure 3.2).

Figure 3.2. Index of enquiry-based science teaching, reported by teachers

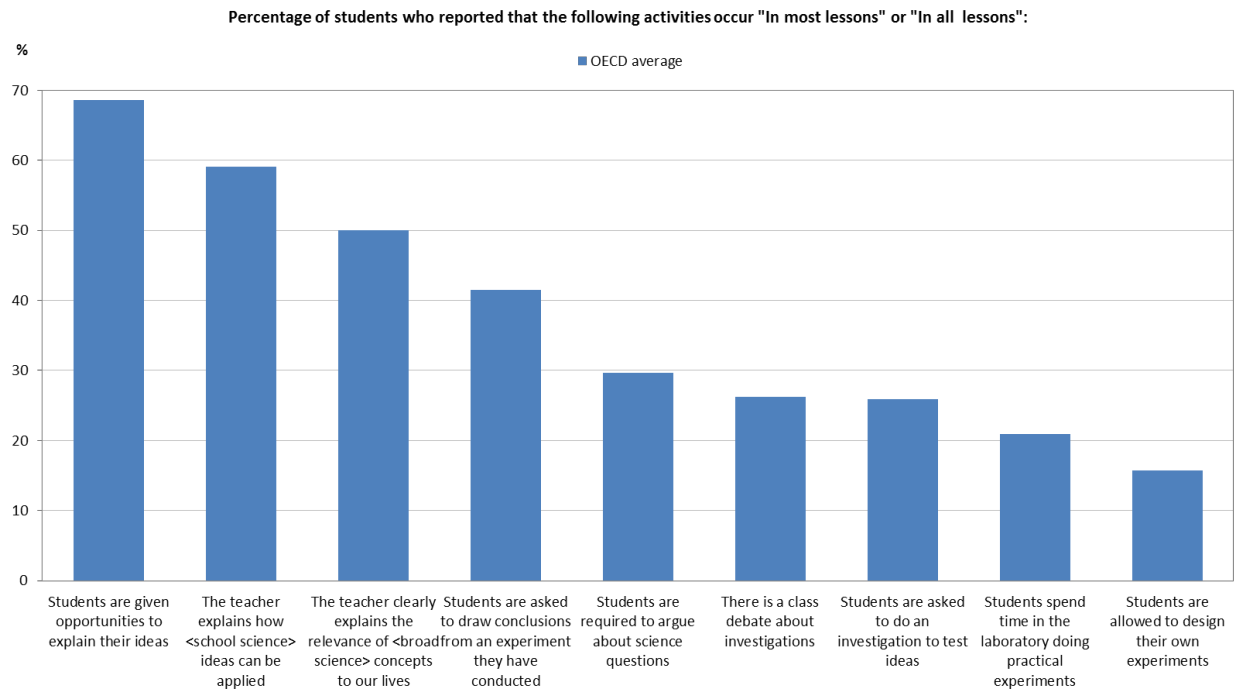


Countries and economies are ranked in descending order of the index of enquiry-based science teaching, reported by students.
Source: OECD, PISA 2015 Database.

Figure 3.3 presents the percentage of students answering "In all lessons" and "In most lessons" to questions about the frequency with which they participate in certain EBST activities in science classes. Across OECD countries, the results show that the most frequent activity students engage in is explaining science ideas (69% of students reported that this happens in all or most lessons). This is followed by the teacher explaining how science ideas can be applied to other phenomena (59% of students so reported), the teacher explaining the broad relevance of science to students' lives (50%), students drawing conclusions from experiments (42%), students arguing about science questions (30%), students debating science investigations in the classroom (26%), students doing an investigation to test ideas (26%), students spending time in the laboratory doing practical experiments (21%), and students designing their own experiments (16%).

The most common activities are those that can be carried out in the classroom and involve discussions between teachers and students. The least common activities are those involving practical work in a laboratory, particularly the one relying on pure discovery, with students designing their own experiments. The reason behind this finding could be that the former group of practices requires fewer resources, is easier to implement and, arguably, could lead to higher achievement in science.

Figure 3.3. Enquiry-based teaching in science lessons



Source: OECD, PISA 2015 Database.

3.2. Enquiry-based science teaching and students' attitudes and expectations

In recent years, more attention has been given to the affective aspects of learning science. Students' current and future engagement with science is shaped by their attitudes and whether they perceive science as important, enjoyable and useful (OECD, 2008). Such attitudes and expectations for the future are likely to be influenced by teaching practices.

In 2015, PISA examined students' enjoyment of science, interest in broad science topics, epistemic beliefs, and their expectation of working in a science-related career when they are 30 years old. This sub-section explores the relationship between EBST and students' attitudes and career expectations.

Enjoyment of science

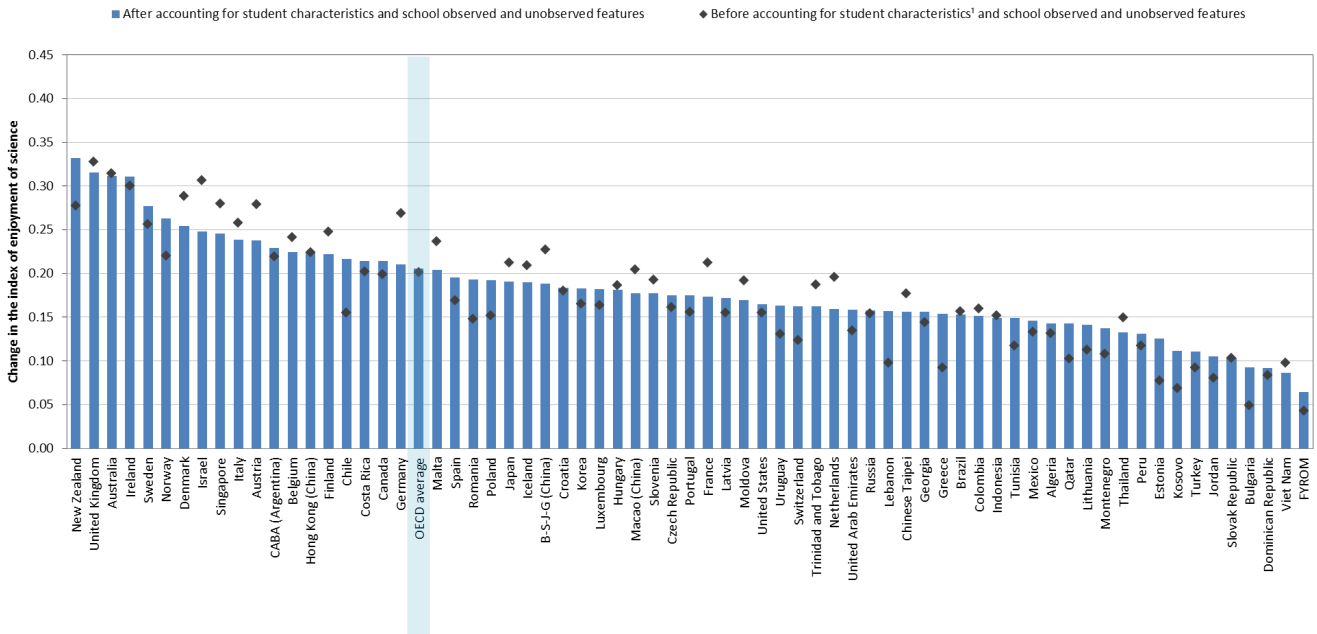
Figure 3.4 presents the results of two regression analyses assessing the association between students' enjoyment of science and reported exposure to EBST. The first analysis relies on a simple regression that does not account for any variable other than EBST; the second relies on a school fixed-effects approach that also takes into account students' socio-

economic status, gender, grade, science performance and the number of science subjects taken.

The results show a positive and significant association between enjoyment of science and student-reported exposure to EBST. The results are moderate in magnitude but statistically significant in all countries and economies. The findings remain significant when observed and unobserved differences among schools and student profiles are accounted for. The associations are modest (regression coefficients ranging between 0.2 and 0.3 in the fixed-effects regression) in Australia, Austria, Belgium, CABA (Argentina), Canada, Chile, Costa Rica, Denmark, Finland, Germany, Hong Kong (China), Ireland, Israel, Italy, Malta, New Zealand, Norway, Singapore, Sweden and the United Kingdom. The weakest associations (regression coefficients smaller than 0.1) are observed in Bulgaria, the Dominican Republic, FYROM, and Viet Nam.

Figure 3.4. Enjoyment of science and enquiry-based science teaching

Change in the index of enjoyment of science associated with a one-unit increase in the index of enquiry-based science teaching, before and after accounting for student characteristics and school observed and unobserved features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: All values are statistically significant.
Countries and economies are ranked in descending order of the change in the index of enjoyment of science associated with a one-unit increase in the index of enquiry-based science teaching reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

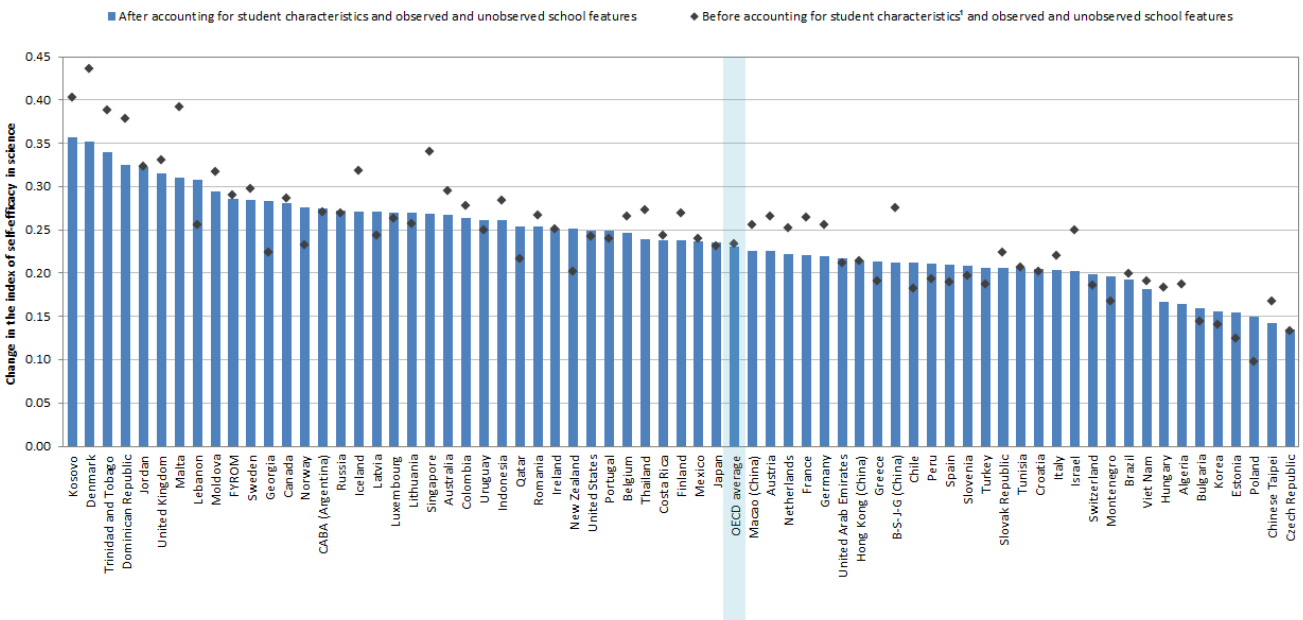
Similar positive but weaker associations are observed between interest in broad science topics and exposure to EBST (Table 3.4). The regression coefficients vary between 0.1 and 0.2 in magnitude for most countries after accounting for school observed and unobserved characteristics and student profile. Australia, Denmark, Ireland, Norway and the United Kingdom stand out as the countries with the largest regression coefficients exceeding 0.2.

Science self-efficacy

Students' self-efficacy in science is defined as students' belief that, through their actions, they can produce desired effects, such as solving a difficult problem or achieving a personal goal. This, in turn, is an incentive to persevere in the face of difficulties (Bandura, 1997). Science self-efficacy refers to one's competency in accomplishing particular goals in a specific context, where meeting these goals requires scientific abilities, such as being able to explain phenomena scientifically, evaluating and designing scientific enquiry, or interpreting data and evidence scientifically (Mason et al., 2013). Such abilities are likely to be affected by science teaching practices that could foster confidence and perseverance when dealing with science-related tasks.

Figure 3.5. Science self-efficacy and enquiry-based science teaching

Change in the index of self-efficacy associated with a one-unit increase in the index of enquiry-based science teaching, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: All values are statistically significant.
Countries and economies are ranked in descending order of the change in the index of self-efficacy associated with a one-unit increase in the index of enquiry-based science teaching reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

The results show that student self-efficacy in science is positively and significantly related to exposure to EBST in all countries, even after accounting for observed and unobserved school characteristics and student profile (Figure 3.5). The magnitude of the regression coefficients varies between 0.2 and 0.3 in most countries and exceeds 0.3 in Denmark, the Dominican Republic, Jordan, Kosovo, Lebanon, Malta, Trinidad and Tobago, and the United Kingdom. In contrast, the associations are particularly weak (below 0.15) in the Czech Republic, Poland and Chinese Taipei.

Science epistemic beliefs

Science literacy as defined in PISA encompasses not only knowledge of science content but also knowledge of how scientific ideas are produced, and an understanding of the goals of scientific enquiry and procedures. In the PISA background questionnaire, students were asked to answer questions about their epistemic beliefs about science (i.e. their beliefs about the nature of knowledge and about the validity of scientific methods as a source of knowledge). Students whose epistemic beliefs are in agreement with current views about the nature of science can be said to value scientific approaches to enquiry (OECD, 2016a).

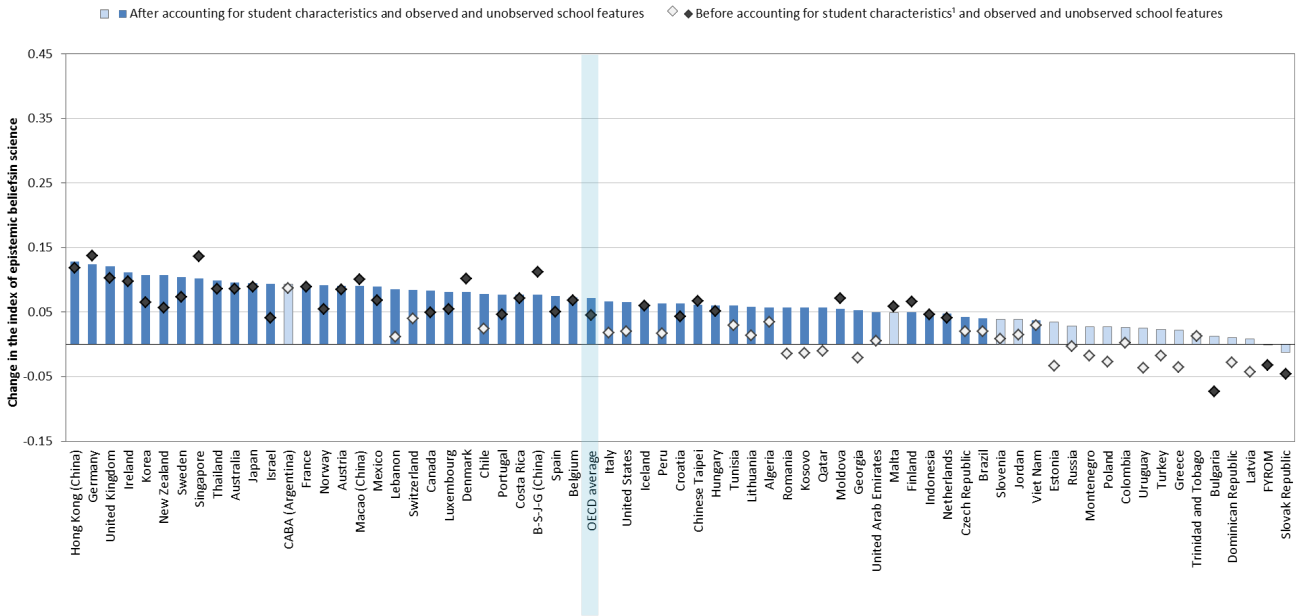
Epistemic beliefs relate to students' understanding about the nature, organisation and sources of scientific knowledge (Hofer and Pintrich, 1997). Students who have a scientific attitude should be able to seek scientific knowledge by adopting an inquisitive approach, searching for data and information, developing scientific hypotheses, testing these hypothesis using robust scientific methods, and finally by interpreting findings and drawing conclusions. Such attitudes and dispositions have been shown to be related both to students' ability to acquire new knowledge in science and to their marks in school science (Mason et al., 2013).

This sub-section examines the relationship between students' epistemic beliefs and their self-reported exposure to EBST. The working assumption is that students exposed to EBST are more likely to develop stronger epistemic beliefs simply because they are experiencing the scientific method first-hand.

Figure 3.6 shows a weak but positive association between students' epistemic beliefs and self-reported exposure to EBST. The regression coefficients range from 0 in FYROM to 0.13 in Hong Kong (China). Positive and significant results are observed in 51 out of 70 countries and economies. Results are statistically non-significant in all other countries and economies. These findings indicate that EBST has the potential to foster students' epistemic beliefs even though the association is weak. The weakness of this association could be related to the demanding nature of successful EBST, which might require the availability of material resources (e.g. laboratory equipment), appropriately trained teachers, and a positive school environment (e.g. discipline in the classroom).

Figure 3.6. Epistemic beliefs in science and enquiry-based science teaching

Change in the index of epistemic beliefs in science associated with a one-unit increase in the index of enquiry-based science teaching, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.

Note: All values are statistically significant.

Countries and economies are ranked in descending order of the likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase in the index of enquiry-based science teaching reported by students, after accounting for student and school characteristics.

Source: OECD, PISA 2015 Database.

Expectations of a science-related career at age 30

As there is growing concern about the capacity of education systems to produce the science skills needed in the labour market, students participating in PISA 2015 were asked about the type of career they expect to be working in when they are 30 years old (OECD, 2016a). The assumption is that children who expect, at an early age, to have a science-related career are more likely to graduate from college or university with a science degree (OECD, 2018; Tai et al., 2006).

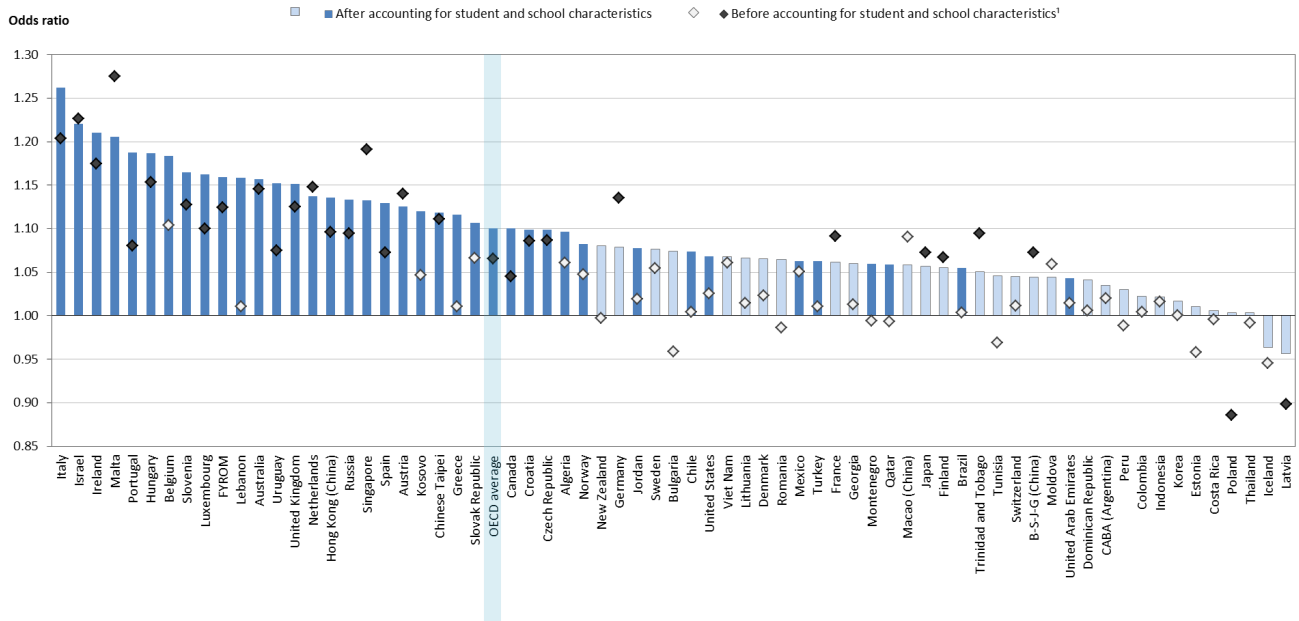
Even though teachers are not professional career educators, their behaviours and teaching practices are expected to influence students' attitudes towards science. Teachers are particularly well-placed to foster students' interest and motivation in science and to cultivate their ambitions (Hutchinson, 2012). Teachers are among the few adults who can help 15-year-old students understand the relationship between learning and earning, and envision their future role in society. Relying on the support of inspired, well-prepared and enthusiastic science teachers, students can understand what their full potential is and, in turn, develop the skills needed to realise their aspirations. In particular, the use of EBST could be the genuine vehicle through which students would experience the excitement of scientific investigation.

Figure 3.7 shows a positive and significant association between student-reported exposure to EBST and the likelihood of expecting a science-related career, even after accounting for student and school profiles. The results are positive and significant in 38 countries and economies. The association is the strongest in Italy, where more exposure to EBST (one additional unit on the index) is associated with a rise of 26% (based on odds ratios) in the likelihood of pursuing a career in science. Italy is followed by Israel (22%), Ireland, Malta (both 21%), Portugal, Hungary (both 19%), and Belgium (18%). In all other countries and economies the association is smaller and non-significant in 30 countries and economies. On average across OECD countries, an increase of one unit on the index of exposure to EBST is associated with a rise of 10% in the likelihood of a student choosing a career in science.

The findings from this section confirm the assumption that exposure to EBST boosts students' enjoyment, interest and self-efficacy in science, and nurtures their epistemic beliefs and science-related career ambitions. The next sub-section explores the relationship between exposure to EBST and actual performance in science.

Figure 3.7. Expectations of a science-related career and enquiry-based science teaching

Likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase in the index of enquiry-based science teaching, before and after accounting for student and school characteristics



1. Student and school characteristics include gender, science performance, the number of science courses attended, the grade in which the student is enrolled, and socio-economic status (as measured by the PISA index of economic, social and cultural status) at the student and school levels.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the change in the index of epistemic beliefs in science associated with a one-unit increase in the index of enquiry-based science teaching reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

3.3. Enquiry-based science teaching and performance in science

One of the major assumptions of the proponents of EBST is that the use of enquiry-based teaching methods would lead to higher achievement in science. In particular, EBST is supposed to familiarise students not only with science content but also with the scientific method and procedures. In other words, EBST consists of an immersion experience through which students get a taste of what it is like to be a scientist. However, as noted in the introduction to this section, the success of EBST hinges on several conditions. For instance, teachers should have the capacity to deliver EBST, which requires time, training and more resources (e.g. laboratory equipment and personnel). The classroom environment should be favourable to the implementation of EBST (e.g. better disciplinary climate). EBST activities also should be adequately designed to cover the appropriate content. The successful design and implementation of EBST activities could compensate for students' lack of experience with scientific methods.

PISA is well positioned to assess the effectiveness of EBST in improving science achievement because, on the one hand, it offers a wealth of information on student and school profiles and, on the other, it assess science literacy rather than the mastery of a particular curriculum. Science literacy in this context refers to students' capacity to apply their knowledge and skills, and to analyse, reason and communicate effectively as they identify, interpret and solve problems in a variety of situations. These are the skills that EBST is expected to foster.

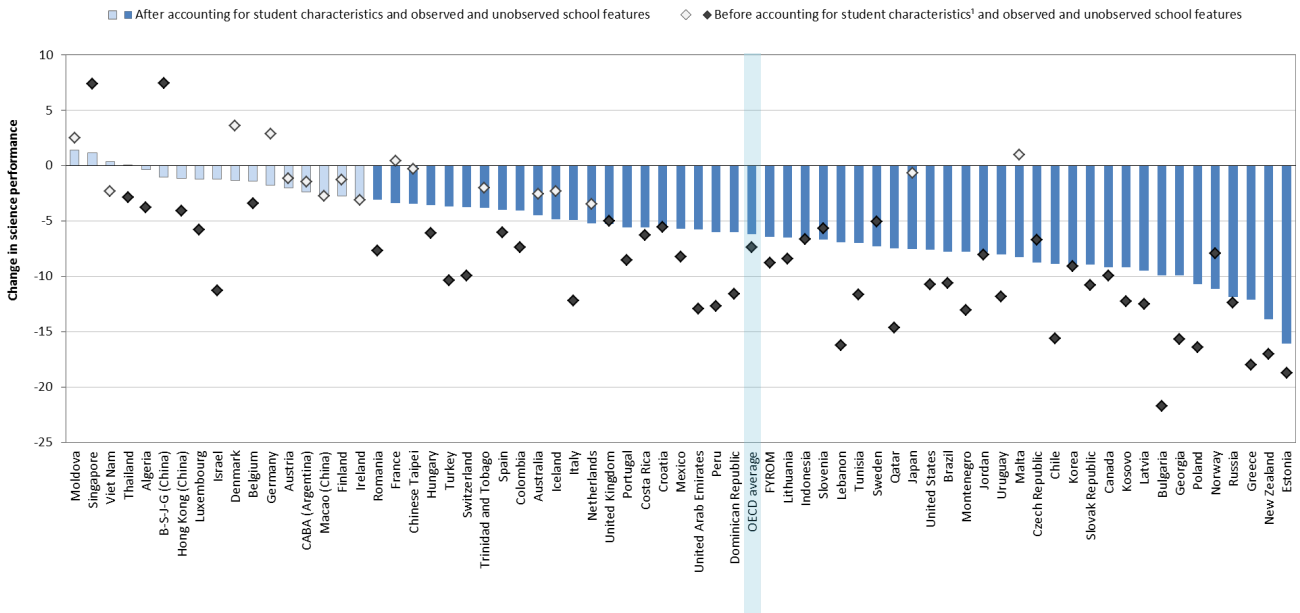
Performance in science and in science subscales

The results presented in Figure 3.8 show a negative association between student-reported exposure to EBST and science performance. The regression coefficients are attenuated in most countries and economies when school observed and unobserved characteristics and student profiles are accounted for; however, they remain negative and significant, albeit weak, in 51 out of 68 countries. This result confirms the findings described in Volume II of the PISA 2015 Report and highlights the need for further investigation.

One possible explanation for this negative association is that students in vocational schools or tracks could be more exposed to EBST than those in academic programmes. This could, in fact, explain the attenuation of the regression coefficients once observed and unobserved school characteristics and student profile have been accounted for. However, the effect remains negative and significant in most countries, which suggests that there is more to this association than meets the eye.

Figure 3.8. Performance in science and enquiry-based science teaching

Change in science performance associated with a one-unit increase in the index of enquiry-based science teaching, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, the number of science courses attended and the grade in which the student is enrolled.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase in the index of enquiry-based science teaching reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

The same regressions were repeated for all science subscales: science competencies, knowledge categories, and content area (the subscales are discussed in detail in Annex A). The findings are mostly identical to the regressions carried out using the overall science scores. This finding indicates that the negative association between exposure to EBST and science performance is not determined by the type of scientific competency, scientific process and knowledge or even by science subjects (e.g. physics, biology, earth sciences) being assessed. As such, the explanation of the negative association lies elsewhere.[3]

Box 3.2. Teacher-reported use of EBST and student outcomes

Similar analyses were carried out using the teacher-reported index of use of EBST (Tables 3.12 to 3.17). The school average of teachers' use of EBST is positively, but weakly, associated with students' enjoyment of science (5 out of 18 countries), students' interest in science (4 out of 18 countries), students' self-efficacy in science (4 out of 18 countries) and, to a lesser extent, with students' epistemic beliefs (1 out of 18 countries) and with students' expectations of a career in science at age 30 (2 out of 18 countries). The associations are weak and attenuated when students' and schools' profiles are accounted for.

When it comes to science scores, the associations with teacher-reported use of EBST are positive and moderate. The associations are statistically significant in 11 out of 18 countries and economies before accounting for students' and schools' profiles, and in 3, albeit weak in magnitude, after accounting for those variables.

Classroom climate and enquiry-based science teaching

Even though EBST might foster enjoyment and self-efficacy in science, these positive associations fail to translate into higher science achievement. A possible explanation is that EBST is a demanding teaching strategy that requires particular resources and school climates. This sub-section examines the interaction between exposure to EBST and student and school contexts. In particular, it focuses on student-reported interest, enjoyment and motivation in science, and on students' sense of belonging at school, perception of teacher fairness and discipline in science classes; and on information reported by school principals on shortages of resources and on student- and teacher-related behavioural problems affecting instruction.

In addition to school observed and unobserved characteristics, the analysis included student profile and student-reported exposure to EBST, and the interaction between EBST and student and school variables (presented in the previous paragraph), which were introduced as quartiles. In other words, students in each country were classified into four quarters on these variables, with each quarter containing 25% of the students.

The results show that the association between EBST and performance varies according to students' perception of discipline in science classes and according to their sense of belonging at school, but not according to the other aforementioned variables. Figure 3.9 presents the science score point change for students in the top and bottom quarters of discipline in science classes who are exposed to EBST (i.e. those students who reported the most disciplined science classes). The score point change was computed using a regression taking into account exposure to EBST (one additional unit on the index) and the four quarters of the index of discipline in science classes with the bottom quarter being the reference category.

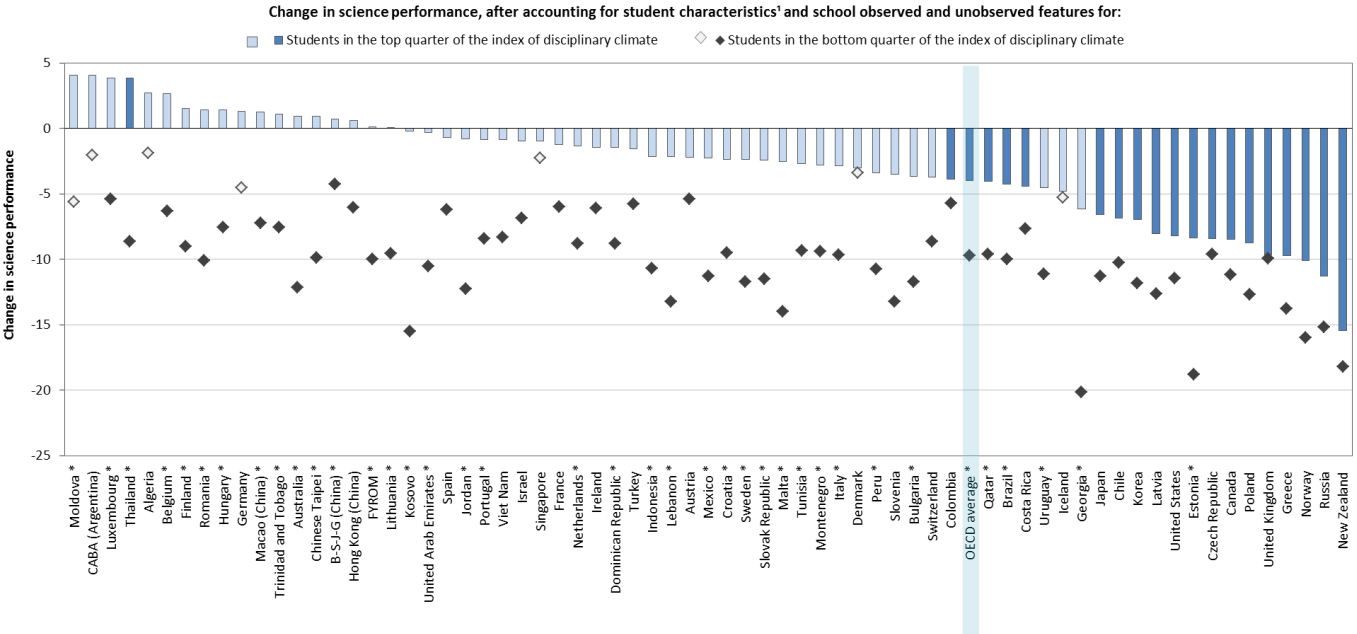
The figure shows that students in the bottom quarter of discipline in science classes perform worse when exposed to EBST in almost all countries and economies. In contrast, students in the top quarter of discipline perform better when exposed to EBST (in comparison with those in the bottom quarter) in 36 countries and economies out of 68 (i.e. in these countries and economies indicated with an asterisk in the figure, the interaction term between EBST and top quarter of discipline is statistically significant). In particular the negative

association between EBST and science performance disappears (in 33 countries and economies) if students are learning in a disciplined environment. In this case the positive interaction between EBST and discipline is positive and strong enough to over-compensate for the negative direct effect of EBST.

In Thailand, students in the top quarter of discipline in science classes derive a positive increase in performance from exposure to EBST (about 4 performance score points) in comparison with their peers in the bottom quarter who underperform by about 13 points. The benefits of attending a disciplined science lesson while being exposed to EBST (as opposed to attending a lesson with a negative disciplinary climate) are the largest in Georgia (20 points), Kosovo (15 points), Lebanon (13 points), Malta (14 points), and Slovenia (13 points).

Figure 3.9. Performance in science and enquiry-based science teaching, by disciplinary climate

Change in science performance associated with a one-unit increase in the index of enquiry-based science teaching, by disciplinary climate in science classes



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, the number of science courses attended and the grade in which the student is enrolled.
Note: Statistically significant values are marked in a darker tone.
Statistically significant interactions between the index of enquiry-based science teaching and the top quarter of the index of disciplinary climate in science classes are marked with an asterisk next to the country/economy name.
Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase in the index of enquiry-based science teaching reported by students, for students in the top quarter of the index of disciplinary climate.
Source: OECD, PISA 2015 Database.

A positive and significant interaction was also observed between EBST and students' sense of belonging at school (i.e. students in the top quarter of sense of belonging) in 17 out of 68 participating countries and economies (Table 3.9). Students who have a stronger sense

of belonging tend to benefit more from exposure to EBST compared with those in the bottom quarter of the index of sense of belonging at school.

This positive association between the interaction of EBST, class discipline and sense of belonging at school, on the one hand, and science performance, on the other, indicates that the success of EBST is conditional on a positive school environment in some countries. This finding confirms the assumptions that EBST is a demanding practice that requires certain preconditions. Moreover, other factors, like school resources and student interest and motivation, do not affect the association between EBST and science performance.

Box 3.3. Using technology to support enquiry-based science teaching

Information and communication technologies (ICTs) can provide teachers with the tools to support learning and can help students acquire the digital skills needed for the 21st century. However, the evidence of the effects of digital technologies in the classroom is not conclusive, especially when ICTs are combined with particular teaching practices (Bulman and Fairlie 2016; Falck et al 2015; Rodrigues and Biagi 2017).

In PISA 2015 students were asked to report on the availability and use of ICTs at school (questions IC009 and IC011). Answers to the two questions were then combined to construct two continuous composite indicators. These indicators were then examined in much the same way as the interaction between school climate and enquiry-based science teaching (EBST) was analysed. In addition to observed and unobserved school characteristics, the regressions included student profile and student-reported exposure to EBST, and the interaction between EBST and the two ICT indices.

The results of the regressions show the expected negative association between EBST and science performance. The interaction between EBST and availability of ICT resources in school is non-significant in almost all countries (except Bulgaria and Lithuania, where it is positive and significant, but weak). However, the interaction between EBST and the use of ICT resources is positive and significant in eight countries: Brazil, Bulgaria, the Dominican Republic, France, Lithuania, Poland, Slovenia and Uruguay. In all of these countries the association is weak, and varies between a 5 and 10 score-point change. Results are provided in the tables in the Annex.

There is no clear and overwhelming evidence that EBST would be positively associated with science performance if ICTs are available and used at school to support this teaching practice. The results also suggest that in a few countries, using ICT resources to support learning is more important than just their availability.

Does the relationship between EBST and performance vary according to the levels of student performance?

Further analyses were carried out in order to ascertain whether the pattern of association between student-reported exposure to EBST and overall science performance varies according to the level of performance. The aim is to determine whether EBST is more or less beneficial to high and low performers (e.g. whether low performers benefit more from exposure to EBST in comparison with high performers because EBST makes science concepts easier to understand). This analysis was carried out using quantile regressions as described in section 2, with associations between EBST and performance computed at the 10th, 25th, 50th, 75th and 90th percentiles of performance. The results show that

differences between regression coefficients at the 10th and 90th percentiles are statistically non-significant in all but four countries.

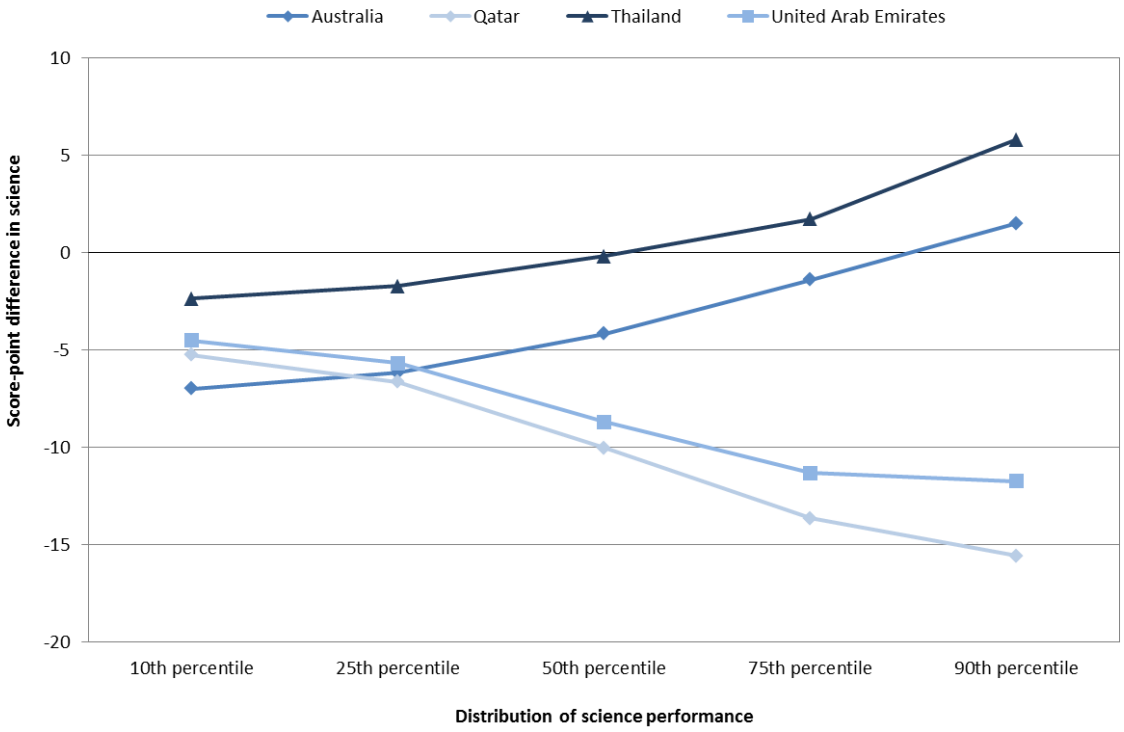
In Australia and Thailand students in the 90th percentile score about 8 points higher than those in the 10th percentile when exposed to one additional unit of EBST (Figure 3.10). In Australia, a one-unit change in EBST exposure is associated with a performance difference of -7 points for students in the 10th percentile of the performance distribution, and with a difference of 1 point for those in the 90th percentile. In Thailand, a one-unit change in exposure to EBST is associated with a performance difference of -2 points for students in the 10th percentile and with a difference of 6 points for those in the 90th percentile.

In contrast, in Qatar and the United Arab Emirates, student in the 90th percentile score about 10 and 7 points lower, respectively, than students in the 10th percentile when exposed to one additional unit of EBST. In Qatar, a one-unit change in EBST exposure is associated with a performance difference of -6 points for students in the 10th percentile of the performance distribution, and with a difference of -16 points for those in the 90th percentile. In the United Arab Emirates, a one-unit change in EBST is associated with a performance difference of -5 points for students in the 10th percentile and with a difference of -12 points for those in the 90th percentile.

The results show that in Qatar and the United Arab Emirates the associations are negative regardless of the level of student performance. The only country where performance seems to be positively associated with exposure to EBST is Thailand, and only for students in the 90th percentile of performance.

Figure 3.10. Science performance and index of enquiry-based science teaching, by performance in science

Score-point difference in science performance associated with a one-unit increase in the index of enquiry-based science teaching (OECD average)



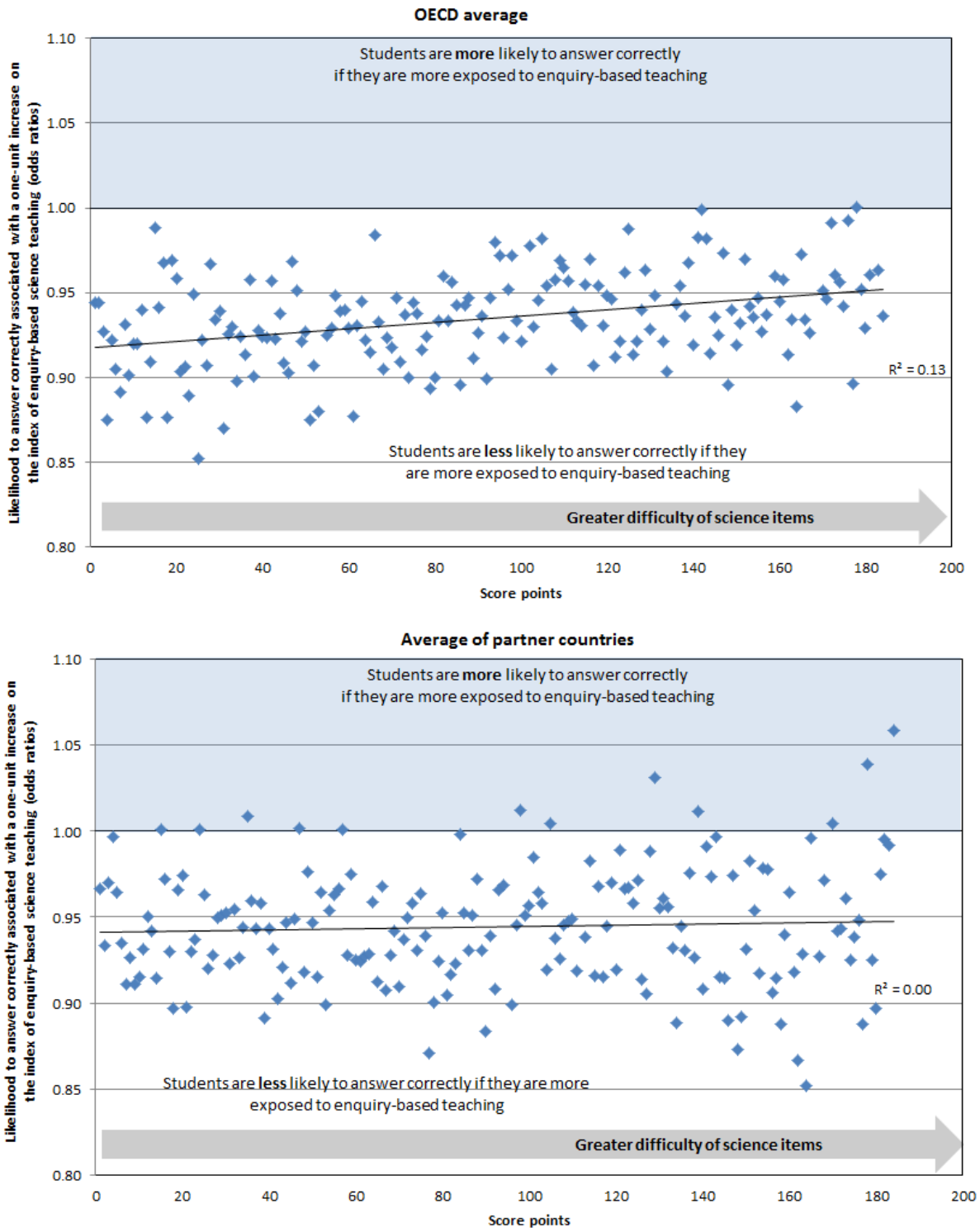
Source: OECD, PISA 2015 Database.

Does EBST help students engage with difficult science tasks?

In order to assess science performance, PISA 2015 used 184 science test items of varying levels of difficulty. Students answered different test items, and the design made it possible to compare the performance of each student on a continuous scale of achievement, as described in Annex A. This sub-section examines the association between student-reported exposure to EBST and the likelihood of answering each test item correctly. By doing so, it is possible to assess whether EBST helps students engage with difficult science tasks in comparison with easier ones.

Answers to each test item were coded as correct or incorrect (1 for correct and 0 otherwise). A logistic regression was then used to explore the association between EBST and the likelihood of answering science test items correctly. All regressions accounted for students' socio-economic status, gender, grade, number of science subjects attended and schools' socio-economic profile. In Figure 3.11, the horizontal axis represents the level of test item difficulty; the vertical axis represents the likelihood (odds ratios) that students answer the item correctly when exposed to one additional unit of EBST. The figure shows that the association between EBST and the likelihood of answering a test item correctly is negative regardless of the level of item difficulty (odds ratios below 1). However, across OECD countries, the association is less negative for the most difficult test items (among partner countries, the association is the same and the line is almost flat).

Figure 3.11. Enquiry-based science teaching and success in science items



Source: OECD, PISA 2015 Database.

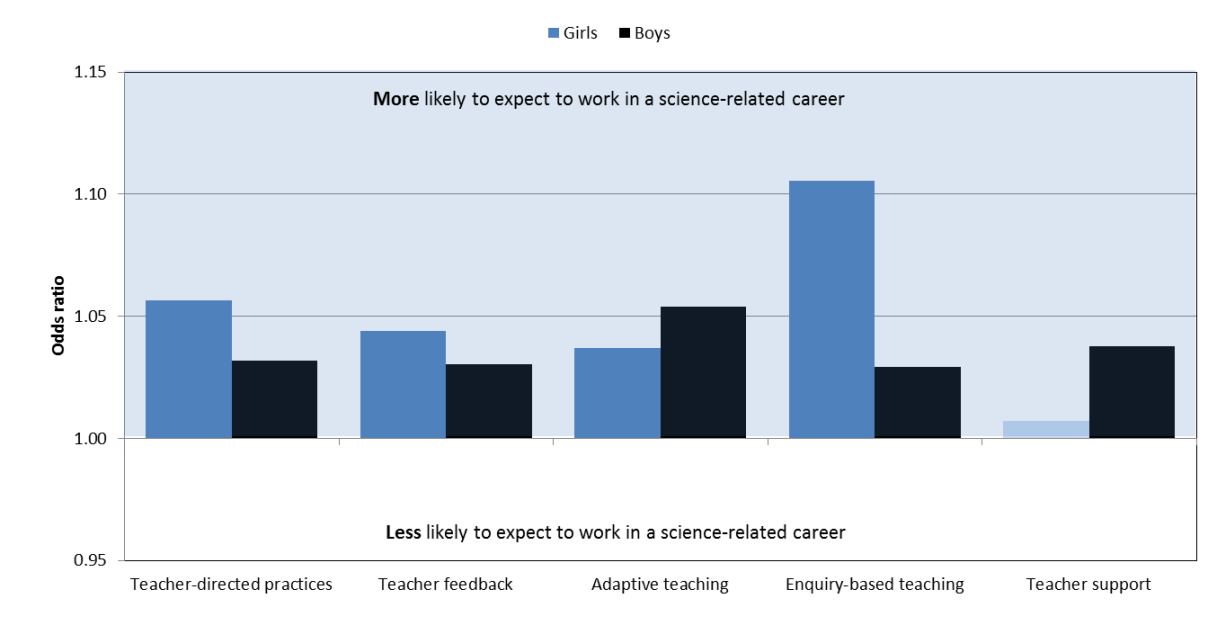
Box 3.4. Getting boys and girls engaged in science

One of the most prominent issues in education today is getting more students, and particularly girls, interested in pursuing a career in science, technology, engineering or mathematics (STEM). In PISA 2015, about 24% of students reported that they expect to work in a science-related career at age 30, on average across OECD countries. However, this proportion may be insufficient to meet labour market needs, given students' declining interest in science as they progress through education (Galton, 2009; Osborne, Simon and Collins, 2003) and the perceived shortage of science graduates in the labour market (Gago et al., 2004; UNESCO, 2016; Olson and Riordan, 2012).

What can science teachers do in their lessons to encourage more students, particularly girls, to consider a career in science? Students' reports about their classes show that, after accounting for science performance, most teaching practices are positively related to students' expectations of pursuing a career in science – especially enquiry-based teaching practices in the case of girls (Figure 3.12). On average across OECD countries, the odds ratio of girls expecting to work in a science-related career increases by about 10% per one-unit increase in the index of enquiry-based teaching. By contrast, boys appear to benefit when their science teachers provide them with greater support.

Figure 3.12. Teaching practices and expectations of pursuing a science-related career

OECD average (34 countries)



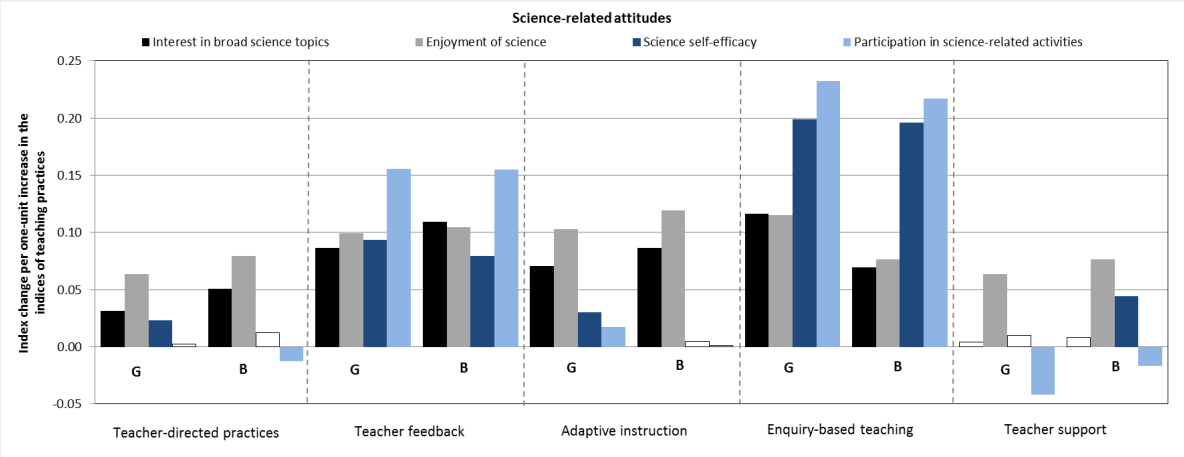
Notes: Results based on logit regression analysis accounting for science performance and other teaching practices.
Statistically significant odds ratios are marked in a darker tone.
Source: OECD, PISA 2015 Database.

To encourage more students to consider a science-related career, teachers can also foster positive attitudes towards science. Again, across OECD countries, exposing students, especially girls, to enquiry-

based practices seems the most promising teaching approach to nurture a series of positive science-related attitudes, including interest in broad science topics, enjoyment of science, science self-efficacy and participation in science-related activities (Figure 3.13). Providing more feedback to students in science lessons appears as the second most promising teaching approach, followed by adaptive teaching and teacher-directed practices.

Figure 3.13. Teaching practices and science-related attitudes

OECD average (34 countries)

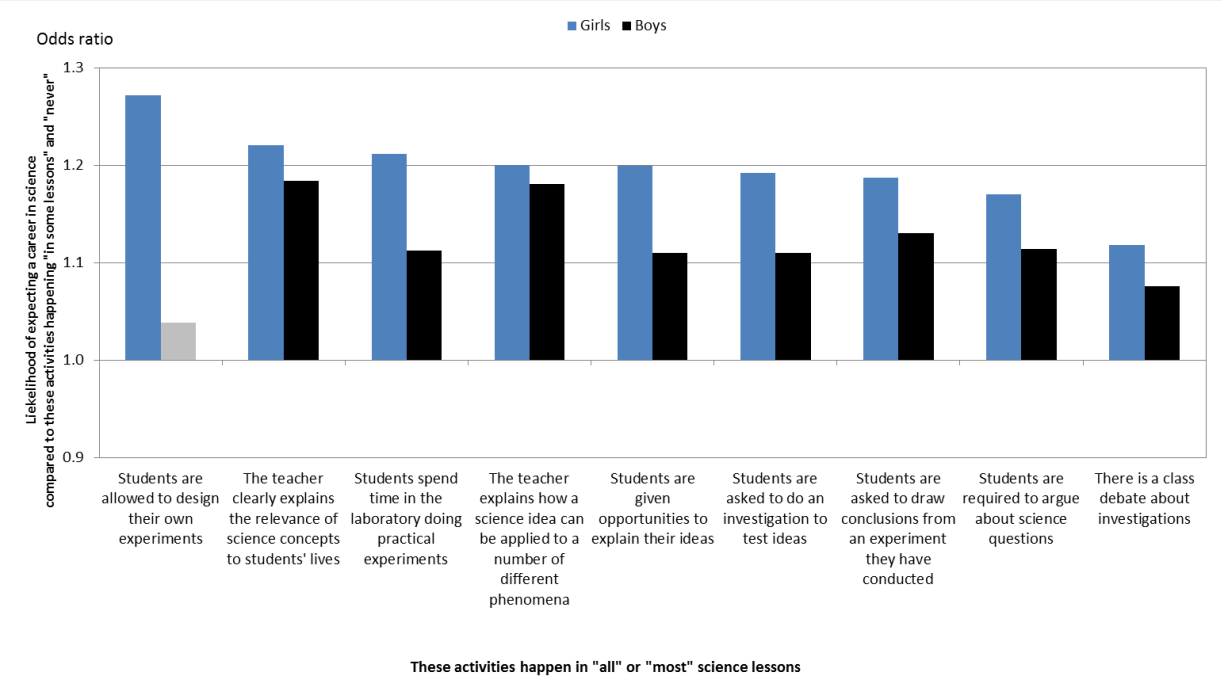


Notes: G: Girls; B: Boys.
Results based on linear regression analyses (one for each science-related attitude) accounting for science performance and other teaching practices.
Statistically not significant coefficients are marked in white.
Source: OECD, PISA 2015 Database.

Among the enquiry-based teaching practices included in the PISA questionnaires, all are positively associated with science-related career expectations (Figure 3.14). Interestingly, all of these practices are more positively associated with girls' expectations of pursuing a career in science, especially when the questions are related to designing experiments and spending time in the laboratory. For instance, the odds ratio of expecting to pursue a career in science is 21% higher among girls who reported that they spend time in the laboratory doing practical experiments in all or most science lessons than among girls who reported that this happened less frequently; the increase among boys was only 11%.

Figure 3.14. Enquiry-based teaching practices and expectations of a science-related career

OECD average



Notes: Statistically significant odds ratios are marked in a darker tone.
Results based on logit regression analysis accounting for science performance.
Source: OECD, PISA 2015 Database.

Despite a similar proportion of girls (24%) and boys (25%) expecting a career in science, boys and girls tend to see themselves working in different fields of science. For instance, only 5% of girls see themselves as engineers, compared to 12% of boys, and only a handful of girls (less than 1%) expect to work as ICT professionals, compared to 5% of boys. Healthcare professions are the main scientific occupations where girls (17%) are more likely to see themselves working in the future than boys (6%), on average across OECD countries. For instance, in 45 education systems, "veterinarians" is among the top five science-related career that girls expect for themselves, compared to only 5 education systems where this is the case for boys. Similar disparities are observed for "nurses and midwives" and "dentists, pharmacists, physiotherapists, dieticians".

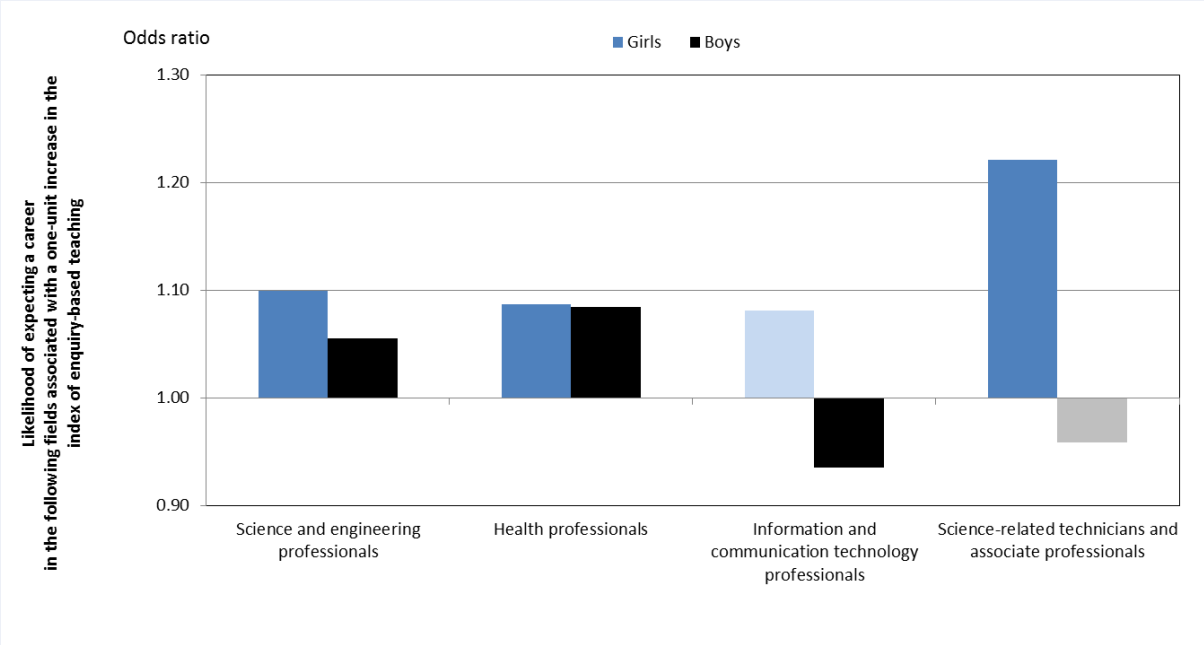
Exposing more girls to enquiry-based teaching practices could not only draw more of them into a science career but may also close the gender gap in traditionally male-dominated science occupations. According to PISA 2015 data, on average across OECD countries, the association between exposure to enquiry-based teaching and expecting a career in science is more positive for girls, but only for male-dominated occupations, including engineers and science-related technicians (Figure 3.15). By contrast, the probability that students see themselves as health professionals – typically female-dominated occupations – increases equally for boys and girls when they are more exposed to enquiry-based teaching.

These findings echo a common recommendation to increase female participation in STEM careers (together with the use of role models and appropriate teacher training): engage girls as early as possible

in investigations, laboratory work and other science-related activities. For instance, the University of Nevada recommends "using online, face-to-face, and other outreach methods to actively recruit girls into voluntary courses and STEM activities";[4] in a policy brief UNESCO proposes to allow (female) "students to apply their learning in real-life situations as well as creative and hands-on experiments" (UNESCO, 2016) and a blog in the *Huffington Post* suggests "purchasing toys and games for girls that encourage problem solving and innovation" and "encouraging group activities and competition (e.g. robotics competition, science fair) among girls that are related to the scientific field".

Figure 3.15. Enquiry-based science teaching and expectation of working in different science-related careers

OECD average (33 countries)



Notes: Statistically significant odds ratios are marked in a darker tone.
Results based on logit regression analysis accounting for science performance.
Source: OECD, PISA 2015 Database.

4. Teacher-directed science instruction

Teacher-directed practices are defined as approaches in which the teacher is largely in control of the content and course of the lesson (Ormrod, 2012). They involve expository instruction – in which information is delivered by the teacher in the form that students are expected to learn it – in addition to class discussions and demonstration of ideas moderated by the teacher. In spite of the common perception that these practices are on the decline (Swaak, de Jong and van Jooligen., 2004), evidence shows that they are still predominant in most countries and are more frequently used than student-oriented strategies (Echazarra et al., 2016; Weiss, 1997).

Teacher-directed instruction often relies on passive learners who are expected to receive information and to reproduce it at a certain point (Cunningham, 1991; Jonassen, 1991). In contrast, enquiry-based learning, as discussed in section 3, relies on the active participation of the learner. The two practices also differ in terms of the learning process. Students exposed to teacher-directed instruction tend, to a certain degree, to rely on learning strategies such as memorisation (Ferguson-Hessler and De Jong, 1990), while those exposed to enquiry-based teaching have to rely on more elaborate strategies that involve collecting, exploring, analysing and interpreting information (de Groot, 1969). By following that process, the learner will be able to infer knowledge (Shute, Glaser and Raghavan, 1989).

The advantages of teacher-directed instruction are orderly classrooms that are easier to manage, wider coverage of subject content, student independence and better preparation for standardised tests (Ormrod, 2012). In contrast, the disadvantages of being exposed solely to this practice are less motivation, negative attitudes towards the subject being learned, and little or no use of collaboration and communication skills.

Enquiry-based methods, on the other hand, have the advantage of fostering communication and collaboration, encouraging students to direct their own learning and developing an authentic interest for a subject. However, these techniques come at the cost of being more challenging for the teacher to implement and being more likely to fail in imparting the required knowledge if no guidance is provided to the students. In fact, in the case of unguided enquiry-based learning, students may develop or infer incorrect knowledge that can be hard to shift once reinforced.

A number of studies provide evidence in favour of teacher-led instruction. In particular, (Shute, 1990) investigated the difference between inductive and deductive learning environments. The authors found a positive effect associated with feedback given by teachers. (Rieber and Parmley, 1995) compared a physics lesson with or without a tutorial and with or without a structure. They found that students who took the unstructured lesson without a prior teacher-provided tutorial performed significantly worse. (Swaak, de Jong and van Joolingen, 2004) compared expository instruction to discovery learning and found that both resulted in substantial learning gains, even though evidence was stronger in favour of the former. (Schwerdt and Wuppermann, 2011) explored the impact of lecture style instruction on students' achievements and found a positive effect that persisted even after accounting for possible selection bias.

In this section PISA data are used to examine the impact of teacher-directed science instruction (hereafter TDSI) on science-related cognitive and non-cognitive outcomes.

Box 4.1. Teacher-directed science instruction

The use of teacher-directed science instruction (TDSI) was measured in PISA using both student- and teacher-reported information.

For the students, one question (question ST103) with 4 items was used. The question asked about the frequency with which certain instructional practices were undertaken by teachers in science classes. Answers were provided on a four-point Likert scale ranging from "Never or almost never", "Some lessons", "Many lessons", to "Every lesson or almost every lesson". The question is:

How often do these things happen in your lessons for this <school science> course?

1. The teacher explains scientific ideas.
2. A whole-class discussion takes place with the teacher.
3. The teacher discusses our questions.
4. The teacher demonstrates an idea.

An index was constructed based on answers to these four statements using IRT scaling (see OECD, 2017c, *Technical Report*, Chapter 16). The index was standardised to have an average of 0 across OECD countries and a standard deviation of 1.

For the teachers, one question (question TC037) with 22 items was used to measure the frequency with which they used certain activities in science classes. Four items correspond to teacher-directed science instruction. Answers were provided on a four-point Likert scale ranging from "Never or almost never", "Some lessons", "Many lessons", to "Every lesson or almost every lesson". The question is:

How often do these things happen in your <school science> lessons?

1. I explain scientific ideas.
2. A whole-class discussion takes place in which I participate.
3. I discuss questions that students ask.
4. I demonstrate an idea.

These four items were used to construct a teacher-reported index of reliance on teacher-directed instruction using the method described in Box 2.1.

For both student- and teacher-reported indices, higher values on the indices indicate more frequent use of teacher-directed instruction.

4.1. Evidence of the use of teacher-directed science instruction

Figure 4.1 presents the average of the student-reported index of exposure to TDSI. Higher values on the index indicate greater student exposure to this teaching practice. The findings show that TDSI is more commonly used in Australia, Canada, Finland, Greece, Iceland, Jordan, Kazakhstan, Lebanon, New Zealand, Poland, Portugal, the Russian Federation (hereafter "Russia"), Singapore, Thailand, Tunisia, the United Arab Emirates and the United States. All of these countries have an average on the index higher than 0.2. In contrast, students are less exposed to TDSI (average lower than -0.2) in Belgium, Costa

Rica, the Czech Republic, Germany, Japan, Korea, Kosovo, Montenegro, the Netherlands, Romania, the Slovak Republic and Uruguay.

Variations in student exposure to TDSI are the largest in Algeria, Bulgaria, Jordan, Kosovo, Lithuania and Qatar, and the smallest in Albania, Ciudad Autónoma de Buenos Aires (Argentina) (hereafter "CABA [Argentina]"), Denmark, Georgia, Hong Kong (China), Indonesia, Italy, Japan, Latvia, Macao (China), Moldova, the Netherlands, Romania, Spain, Uruguay and Viet Nam. Small variations indicate that students have similar levels of exposure to TDSI.

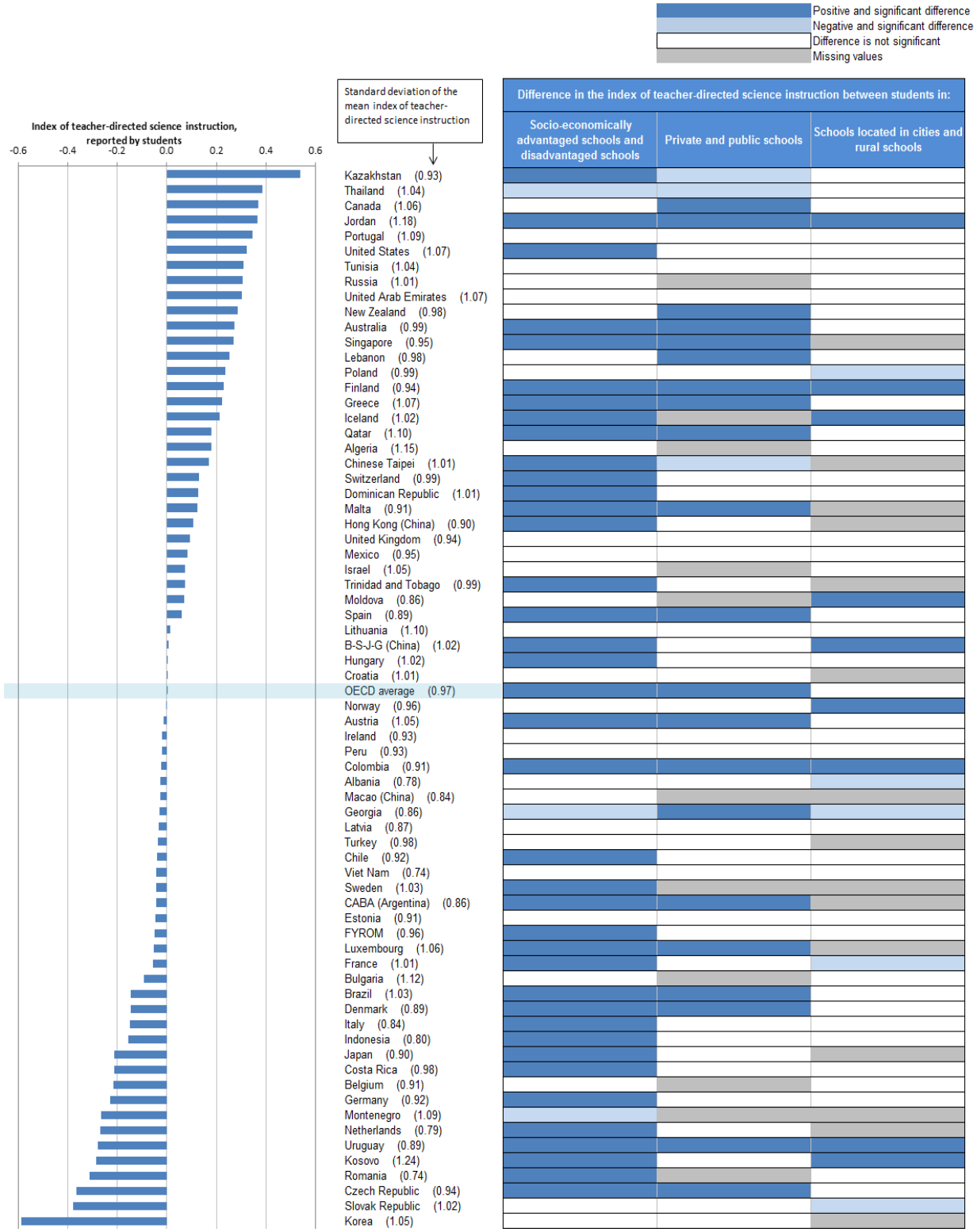
Student exposure to TDSI varies according to the characteristics of their schools. The results show that students in the most advantaged schools (schools in the top quarter of socio-economic status, as measured by the PISA index of economic, social and cultural status) tend to report more frequent exposure to TDSI than those attending schools in the bottom quarter of socio-economic status. The difference is particularly large in Australia, Austria, Beijing-Shanghai-Jiangsu-Guangdong (China) (hereafter "B-S-J-G [China]"), Brazil, CABA (Argentina), Colombia, Costa Rica, Germany, Greece, Hungary, Iceland, Japan, Jordan, Kazakhstan, Kosovo, Luxembourg, Malta, the Netherlands, Romania, Singapore and Sweden. One possible explanation is that TDSI is less frequently used in vocational schools which tend to have lower average ESCS and to rely more on EBST (see previous section).

Similarly, exposure to TDSI is more frequently reported by students in private schools than by students in public schools. The difference is the largest in Brazil, Canada, Colombia, Finland, Georgia, Greece, Malta and Portugal. These findings show the opposite associations in comparison with exposure to enquiry-based science teaching (EBST), which is more frequently reported by students in disadvantaged and in public schools.

When it comes to school location, the results show limited differences in student-reported exposure to TDSI. In B-S-J-G (China), Colombia, Finland, Iceland, Jordan, Kosovo, Norway and Uruguay, students attending urban schools are more exposed to TDSI than students in rural schools; the reverse is true in Albania, France, Georgia and Poland.

Few differences are observed in terms of the profile of the student population in a school (Table 4.2). In Norway, Qatar and Singapore, TDSI is more frequently reported by students in schools where more than 30% of students have an immigrant background (compared with schools where less than 30% of students have an immigrant background). The reverse is true in Germany, Hong Kong (China) and the United States. Moreover, in Canada, Norway and Qatar, students reported greater exposure to TDSI in schools where more than 30% of students speak a different language at home from the language of assessment (compared with schools with less than 30% of such students). The opposite is observed in France, Macao (China), the United Arab Emirates and the United States.

Figure 4.1. Index of teacher-directed science instruction, reported by students

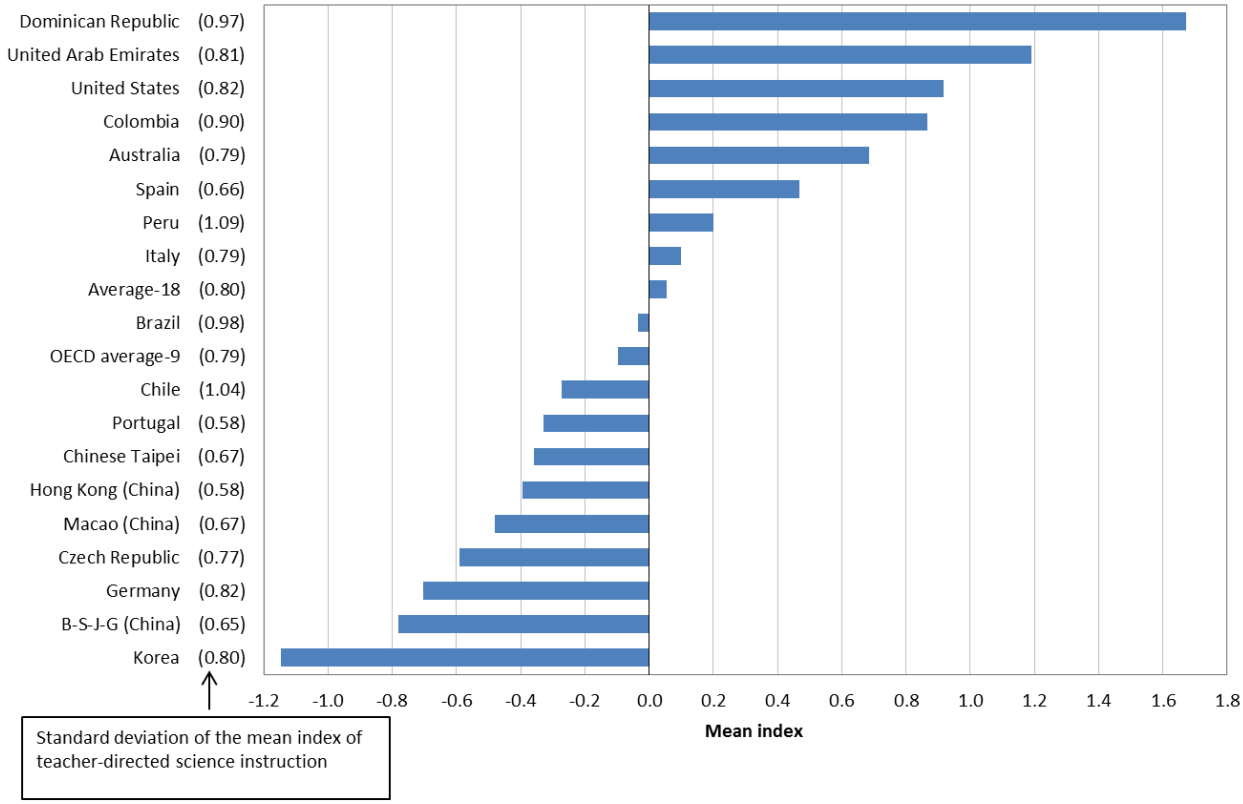


Countries and economies are ranked in descending order of the index of teacher-directed instruction, reported by students.
Source: OECD, PISA 2015 Database.

In general, students' exposure to TDSI is related to the characteristics of their schools. Students are more exposed to TDSI if they attend socio-economically advantaged private schools. The findings also show that teacher-directed instruction is still common across many education systems, including Arabic-speaking and some English-speaking countries, such as Australia, New Zealand and the United States.

As reported by teachers, TDSI is frequently used in Australia, Colombia, the Dominican Republic, Spain, the United Arab Emirates, and the United States. It is less frequently used in B-S-J-G (China), Chile, the Czech Republic, Germany, Hong Kong (China), Korea, Macao (China), Portugal and Chinese Taipei.

Figure 4.2. Index of teacher-directed instruction, reported by teachers



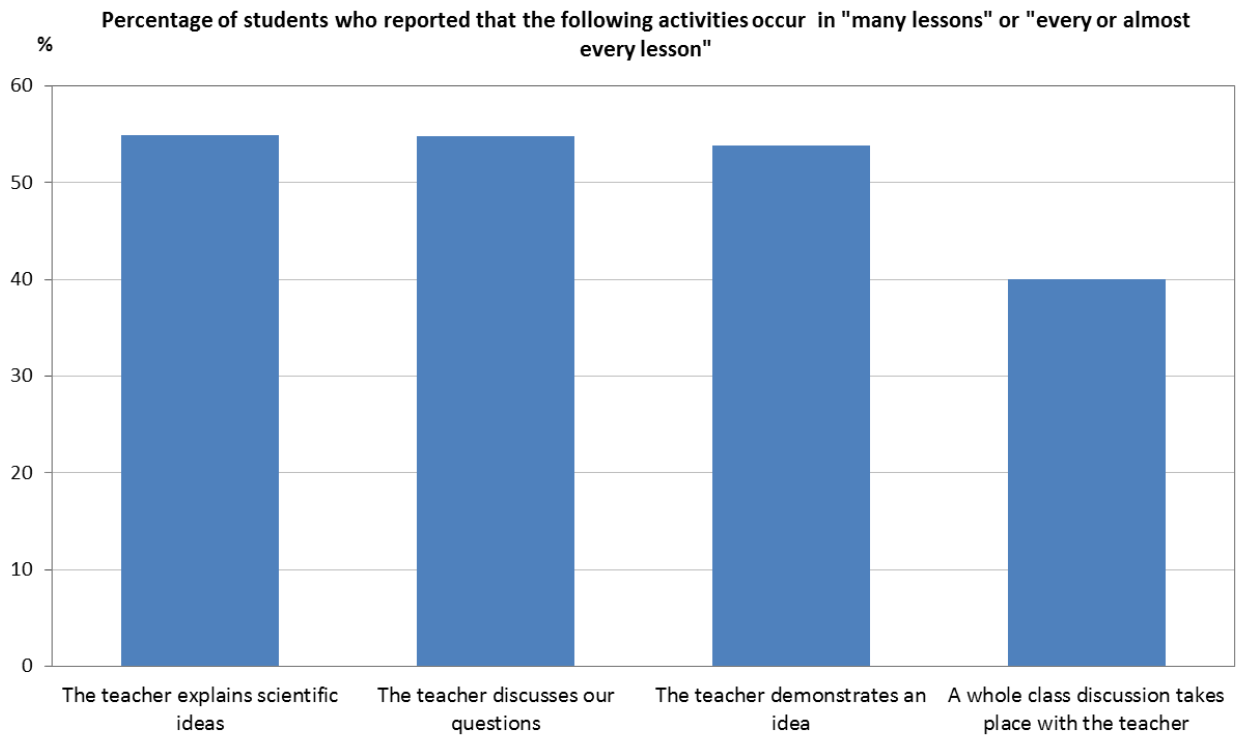
Countries and economies are ranked in descending order of the index of teacher-directed instruction, reported by teachers.
Source: OECD, PISA 2015 Database.

Figure 4.3 presents the proportion of students answering "Many lessons" or "Every lesson or almost every lesson" to questions about the frequency of exposure to TDSI activities. The results show that the most frequent activities reported (by about 54% of students) are: the teacher explains scientific ideas; the teacher discusses students' questions; and the teacher demonstrates ideas. Some 40% of students reported that class discussions with the teacher occur in many, every or almost every lesson. Thus, according to students, the sole

activity involving class discussion occurs less frequently than activities led by teachers, with students as passive learners.

Figure 4.3. Teacher-directed instruction in science lessons

OECD average

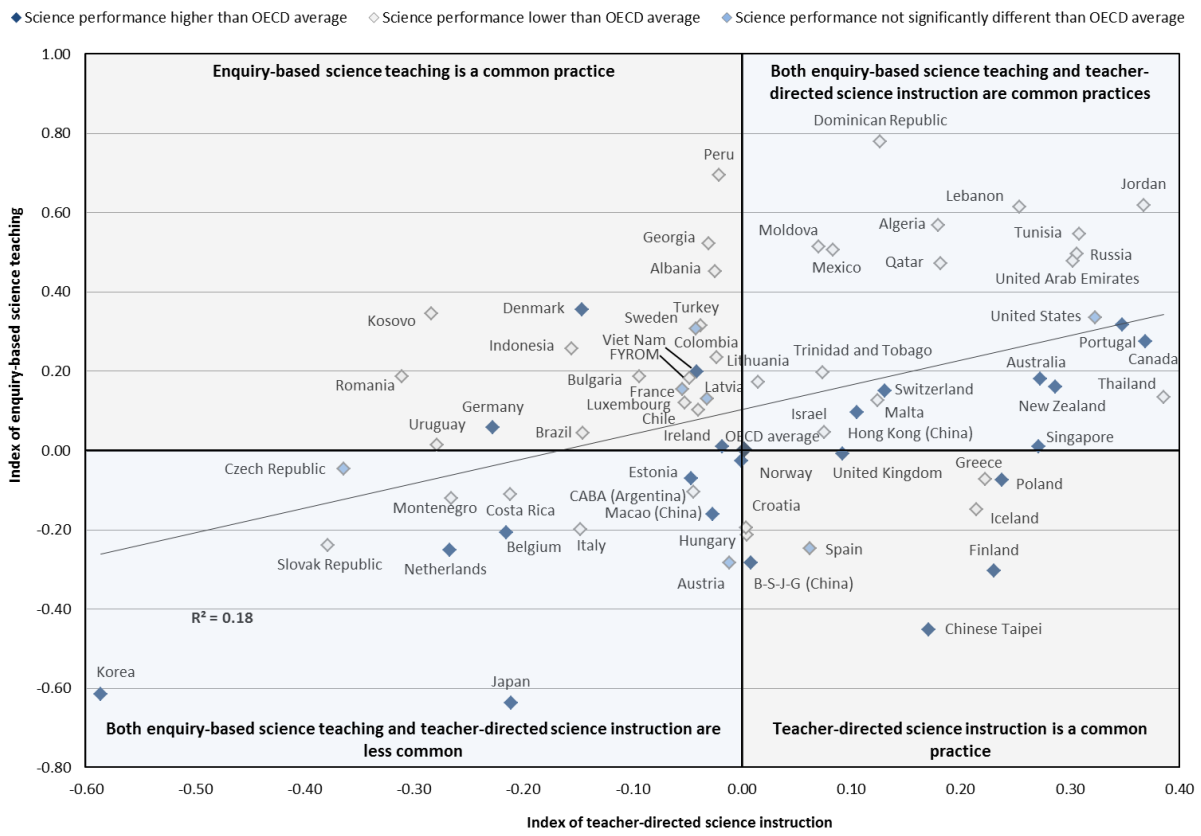


Source: OECD, PISA 2015 Database.

When students' exposure to TDSI is plotted against their reported exposure to EBST, the results show that the two practices are not mutually exclusive (Figure 4.4). Both practices are common in Algeria, Jordan, Lebanon, Portugal, Qatar, Russia, Tunisia, the United Arab Emirates and the United States. The two practices are less common in Belgium, Japan, Korea, Montenegro, the Netherlands, and the Slovak Republic. In Finland and Chinese Taipei, teacher-directed instruction is common but enquiry-based teaching is not. The opposite is true in Denmark, Indonesia, Kosovo and Romania.

From a pedagogical perspective, the joint use of EBST and TDSI could help foster learning in that students are provided with teacher direction and prior knowledge about science topics before engaging in enquiry-based activities. The interaction between the two teaching practices will be investigated further in the next sub-section.

Figure 4.4. Relationship between teacher-directed science instruction and enquiry-based science teaching



Source: OECD, PISA 2015 Database.

4.2. Teacher-directed science instruction and students' attitudes and expectations

This sub-section explores the relationship between students' reported exposure to TDSI and students' non-cognitive outcomes. As in the previous section, the focus is on students' attitudes towards science and whether they perceive science as important and enjoyable. In particular, enjoyment, interest, self-efficacy, epistemic beliefs and career expectations at age 30 are explored.

While a differentiated approach to teaching is seen as a method of catering to the needs of a diverse student population (Gardner, 2011), its limitations include pressure on teachers to develop and adapt their teaching materials and extend instructional time. In contrast, traditional teacher-directed practices have the advantage of being more established and easier to implement; yet they could come at the cost of not being fully suitable to the needs of some students or being less enjoyable than discovery-based practices.

Enjoyment and interest in science

Figure 4.5 presents the results of regression analyses of the association between TDSI and students' enjoyment and interest in science. The first analysis relies on a univariate regression that only takes into account TDSI as an independent variable. The second

accounts for student characteristics (student ESCS, gender, grade, science performance and number of science subjects taken), and overserved and unobserved school features.

The findings show a clear positive and significant association between student-reported exposure to TDSI and enjoyment of science in all countries and economies. The association is moderate (regression coefficients range between 0.2 and 0.3) in 31 countries and economies and weak in all others. It exceeds 0.25 in Australia, CABA (Argentina), Hong Kong (China), Ireland, Israel, the Netherlands, New Zealand, and the United Kingdom.

When it comes to interest in science, the findings are similar even though the associations are weaker. TDSI is positively and significantly related to interest in science in all countries, and the association is moderate in the Dominican Republic and the Netherlands. In all other countries/economies the association is weak. Interestingly, when student characteristics and school observed and unobserved features are accounted for, the association is only slightly modified. In other words, this association is not sensitive to the socio-economic background of the students or to their science performance and school's characteristics and policies.

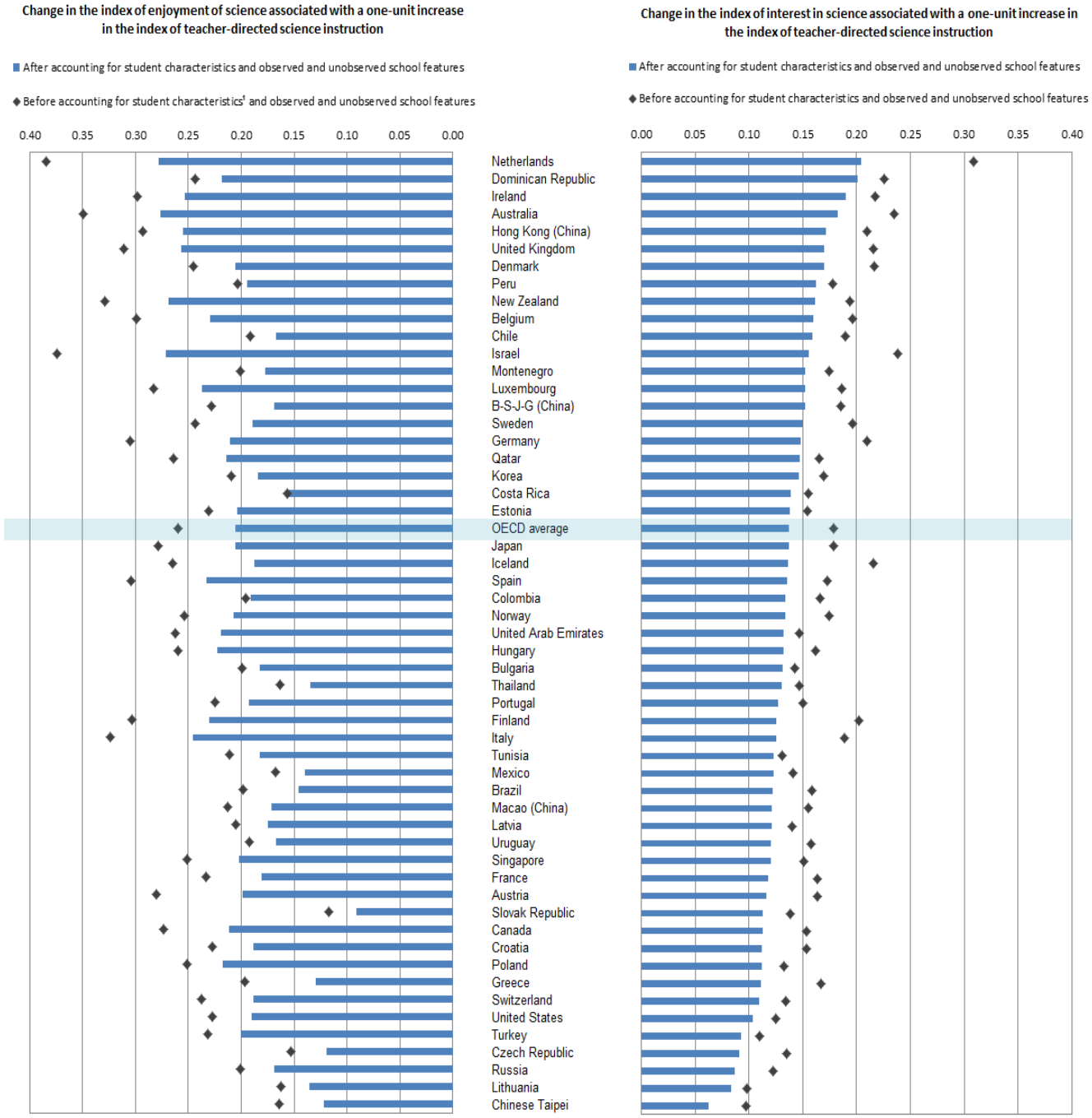
The findings suggest an absence of evidence that students' enjoyment of and interest in science are hindered by exposure to TDSI. On the contrary, the positive associations indicate that TDSI, and teacher-led initiation into science topics more generally, have the potential to foster enjoyment in, interest in and enthusiasm for the subject. And both TDSI and EBST are positively correlated with enjoyment and interest in science. In other words, two different strategies could foster positive attitudes towards science.

Science self-efficacy and epistemic beliefs

Both science self-efficacy and epistemic beliefs are positively and significantly correlated with TDSI, even after accounting for student characteristics and observed and unobserved school features (Figure 4.6). For science self-efficacy, the associations are moderate, with regression coefficients exceeding 0.2 only in the United Kingdom. They are weak in all other countries and are statistically non-significant in Algeria, Bulgaria, CABA (Argentina), the Former Yugoslav Republic of Macedonia (hereafter "FYROM"), Greece, Jordan, Lebanon, Moldova, Montenegro, Romania, the Slovak Republic, Turkey and Uruguay.

Epistemic beliefs in science are significantly, positively but weakly related to TDSI in all countries and economies. The association is lower than 0.1 in B-S-J-G (China), Belgium, Denmark, Greece, Hungary, Indonesia, Kosovo, Latvia, Lithuania, Norway, Portugal, Romania, and the Slovak Republic. The results corroborate those for EBST and show that different teaching strategies could be associated with student attitudes and non-cognitive outcomes.

Figure 4.5. Enjoyment of science, interest in science and teacher-directed science instruction



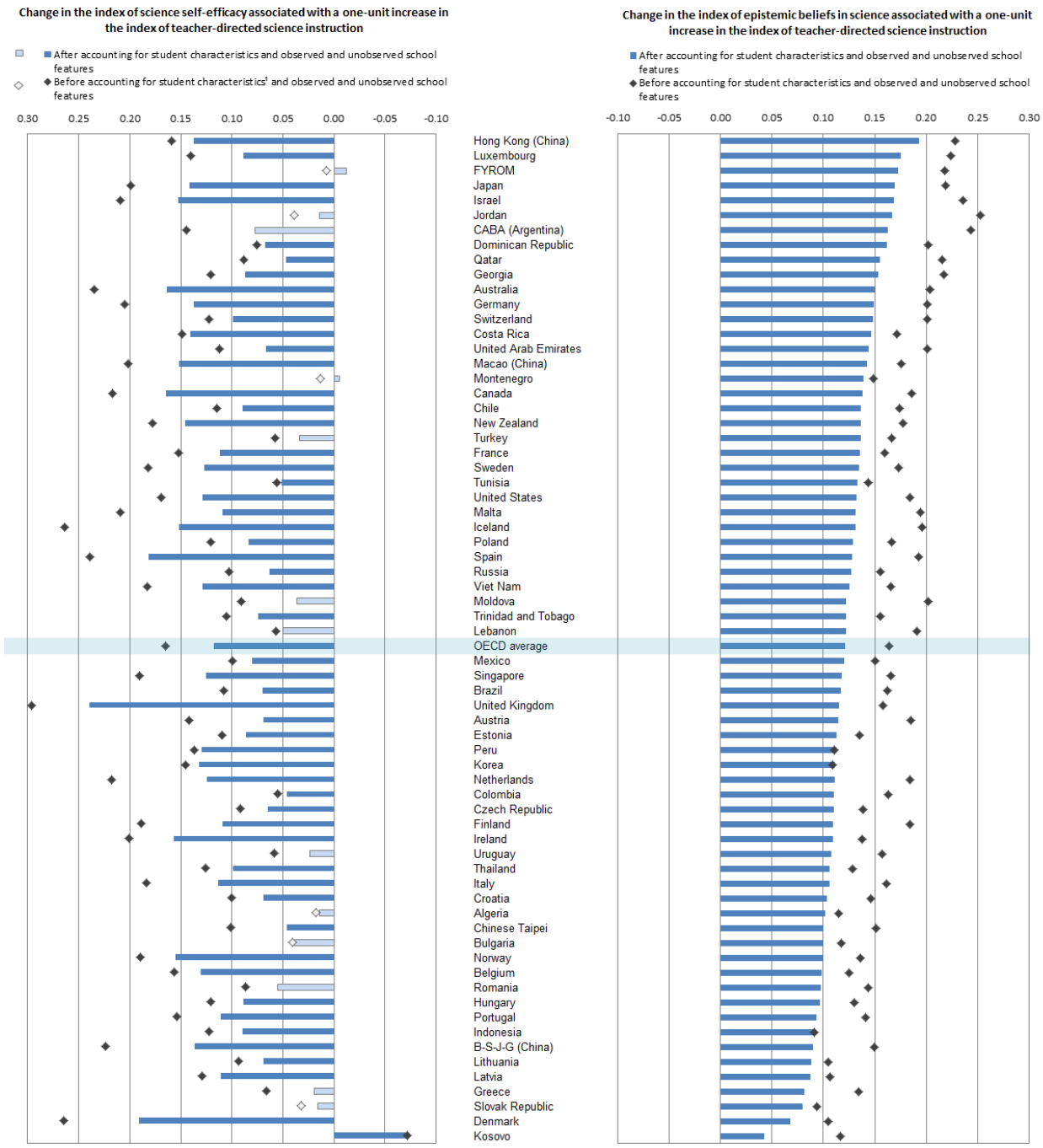
1. Student characteristics include socio-economic status (as measured by the PISA index of economic, cultural and social status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.

Notes: All values are statistically significant.

Countries and economies are ranked in descending order of the change in the index of interest in science associated with a one-unit increase in the index of teacher-directed science instruction, after accounting for student characteristics, and observed and unobserved school features.

Source: OECD, PISA 2015 Database.

Figure 4.6. Epistemic beliefs, science self-efficacy and teacher-directed science instruction



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.

Note: Statistically significant values are marked in a darker tone. All changes in the index of epistemic beliefs in science are significant.

Countries and economies are ranked in descending order of the change in the index of epistemic beliefs in science associated with a one-unit increase in the index of teacher-directed science instruction reported by students, after accounting for student characteristics, and observed and unobserved school features.

Source: OECD, PISA 2015 Database.

Box 4.2. Teacher-reported use of TDSI and student outcomes

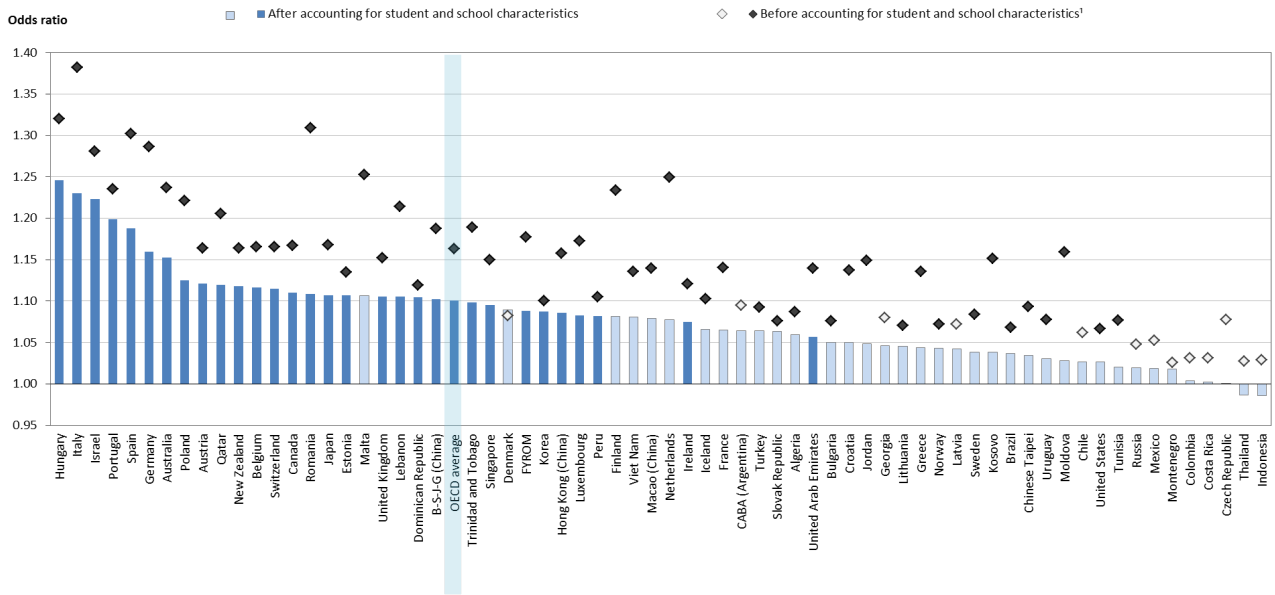
The same analyses were repeated using the teacher-reported index of reliance on TDSI (Tables 3.12 to 3.17). This index was aggregated at the school level and merged with student data due to the absence of a student-teacher link in PISA as explained in section 2. The findings show a positive, statistically significant but weak association between teachers' use of TDSI and students' enjoyment of science (5 out of 18 countries), students' interest in science (3 out of 18 countries), students' science self-efficacy (5 out of 18 countries), students' expectation of a science-related career at age 30 (1 out of 18 countries), and science performance (positive in three countries and negative in 2). Teacher-reported use of TDSI is not significantly associated with students' epistemic beliefs in any of the countries.

Expectations of a science-related career at age 30

The logistic regressions investigating the association between exposure to TDSI and expectation of a science career at age 30 show a positive and statistically significant relationship between the two in 30 out of 67 countries and economies (Figure 4.7). The odds ratio exceeds 1 and is largest in Australia, Germany, Hungary, Israel, Italy, Portugal and Spain. In Hungary, for instance, students who are exposed to one additional unit on the TDSI index are 25% more likely to choose a career in science at age 30. Hungary is followed by Italy (23% more likely), Israel (22%), Portugal (20%), Spain (19%), Germany (16%) and Australia (15%).

Figure 4.7. Expectations of a science-related career and teacher-directed science instruction

Likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase in the index of teacher-directed science instruction, before and after accounting for student and school characteristics



1. Student and school characteristics include gender, science performance, the number of science courses attended, the grade in which the student is enrolled, and socio-economic status (as measured by the PISA index of economic, social and cultural status) at the student and school levels.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the likelihood of students expecting to work in a science-related career at age 30, associated with a one-unit increase in the index of teacher-directed science instruction reported by students, after accounting for student and school characteristics.
Source: OECD, PISA 2015 Database.

4.3. Teacher-directed science instruction and performance in science

This sub-section investigates the relationship between students' exposure to TDSI and their cognitive outcomes. The working assumption is that TDSI could facilitate the transmission of knowledge about science subjects even though it provides less hands-on experience of the procedures of professional scientists. Nonetheless, students might still be able to develop science-related skills, such as critical thinking and problem solving, by discussing science questions with their teachers and peers. PISA's innovative concept of science literacy, which combines knowledge and skills, and the ability to reason and communicate effectively, allows us to go beyond the focus on content knowledge by considering a broader definition of literacy that includes understanding scientific processes and procedures.

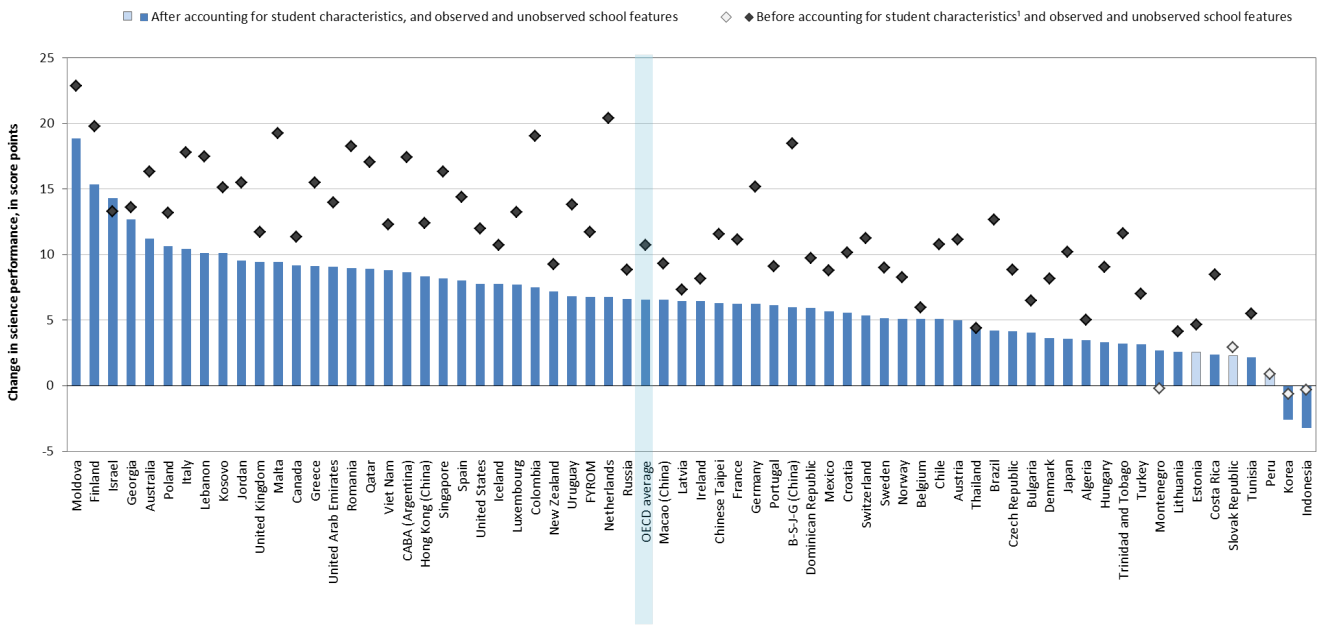
Performance in science and on the science subscales

Figure 4.8 presents the association between exposure to TDSI and overall performance in science. The association is positive in all countries and economies except Indonesia and Korea, where it is negative. It is statistically significant – but weak – in 65 countries and economies even after accounting for student characteristics and observed and unobserved school features. In nine countries, the association exceeds the threshold of a 10-point improvement in science score associated with a rise of one unit on the student-reported

index of TDSI. In Moldova, a rise of one unit on the index is associated with an improvement of 19 score points in science performance. Moldova is followed by Finland (an improvement of 15 points), Israel (14 points), Georgia (13 points), Australia, Poland (both 11 points), Italy, Kosovo and Lebanon (all 10 points). On average across all OECD countries, a one-unit rise on the TDSI index is associated with an improvement of seven points in science performance.[5]

Figure 4.8. Performance in science and teacher-directed science instruction

Change in science performance associated with a one-unit increase in the index of teacher-directed science instruction, before and after accounting for student characteristics, and observed and unobserved school features



1. Student and school characteristics include gender, the number of science courses attended, the grade in which the student is enrolled, and socio-economic status (as measured by the PISA index of economic, social and cultural status) at the student and school levels.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase on the index of teacher-directed science instruction reported by students, after accounting for student characteristics, and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

The same regression analyses were repeated for each of the science subscales: scientific competencies, knowledge categories and content areas. The findings show results similar to those based on the overall science score. The associations are positive, significant but weak in most countries and economies. This finding indicates that the positive relationship between TDSI and science performance does not vary according to the type of scale being assessed. In other words, TDSI is more or less equally associated with both science content knowledge and skills.

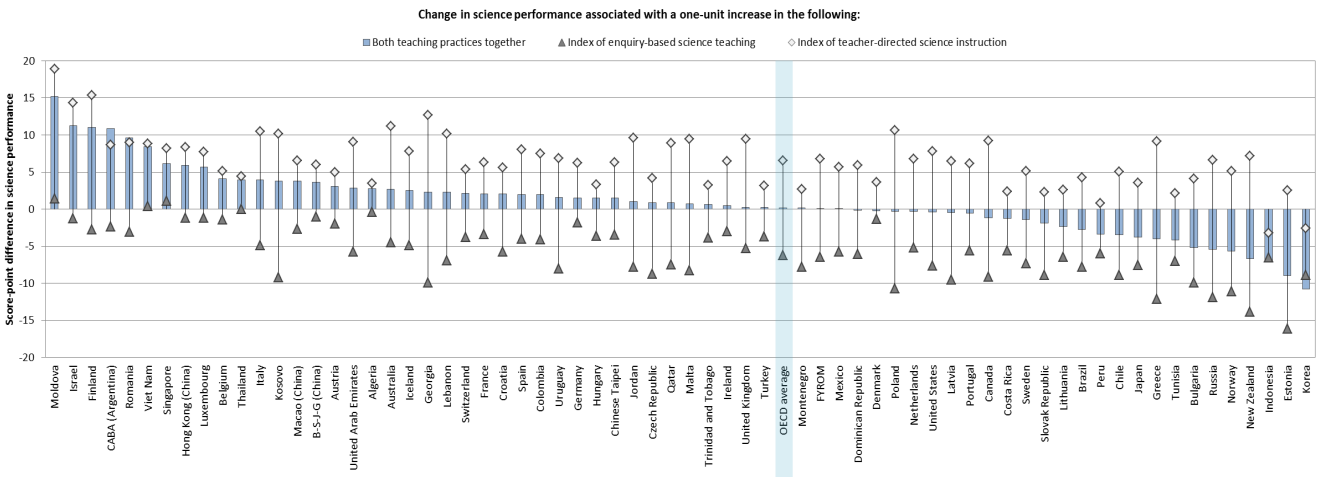
Combining enquiry-based science teaching and teacher-directed instruction

Further analyses were carried out in order to ascertain whether combining EBST and TDSI could help improve student achievement in science. Figure 4.9 presents the results of three regression analyses accounting for teaching practices in addition to student characteristics,

and observed and unobserved school features. The figure shows three statistics: change in science performance associated with a one-unit increase on the index of EBST (triangles); change in science performance associated with a one-unit increase on the index of TDSI (diamonds); and change in science performance associated with a one-unit increase on both indices (bars). The working assumption is that by combining the two practices, teachers will be able to deliver some content knowledge to their students before students engage in enquiry-based activities. As such, students will be equipped with basic knowledge that will guide them on their path to discovery.[6]

As expected, the results show a negative association between science performance and EBST, and a positive association with TDSI when the two practices are taken separately. When the two practices are combined the joint association with science performance is positive in Algeria, Australia, Austria, B-S-J-G (China), Belgium, CABA (Argentina), Colombia, Croatia, the Czech Republic, Finland, France, Georgia, Germany, Hong Kong (China), Hungary, Iceland, Israel, Italy, Jordan, Kosovo, Lebanon, Luxembourg, Macao (China), Malta, Moldova, Qatar, Romania, Singapore, Spain, Switzerland, Chinese Taipei, Thailand, Trinidad and Tobago, United Arab Emirates, Uruguay and Viet Nam. A rise of one unit on both indices of EBST and TDSI is associated with a 15-point improvement in science performance in Moldova and an 11-point improvement in CABA (Argentina), Israel and Finland. But these improvements are not as large as the association with an increase of one unit on the index of TDSI when the practice is used on its own (i.e. when TDSI is not combined with EBST). In other words, the positive relationship between TDSI and science performance is tempered by the concurrent use of EBST.

Figure 4.9. Science teaching practices and student performance in science



Note: Performance is regressed on each index separately, then on both indices taken together and on the interaction between them

Countries and economies are ranked in descending order of the effect of both science teaching practices together.

Source: OECD, PISA 2015 Database.

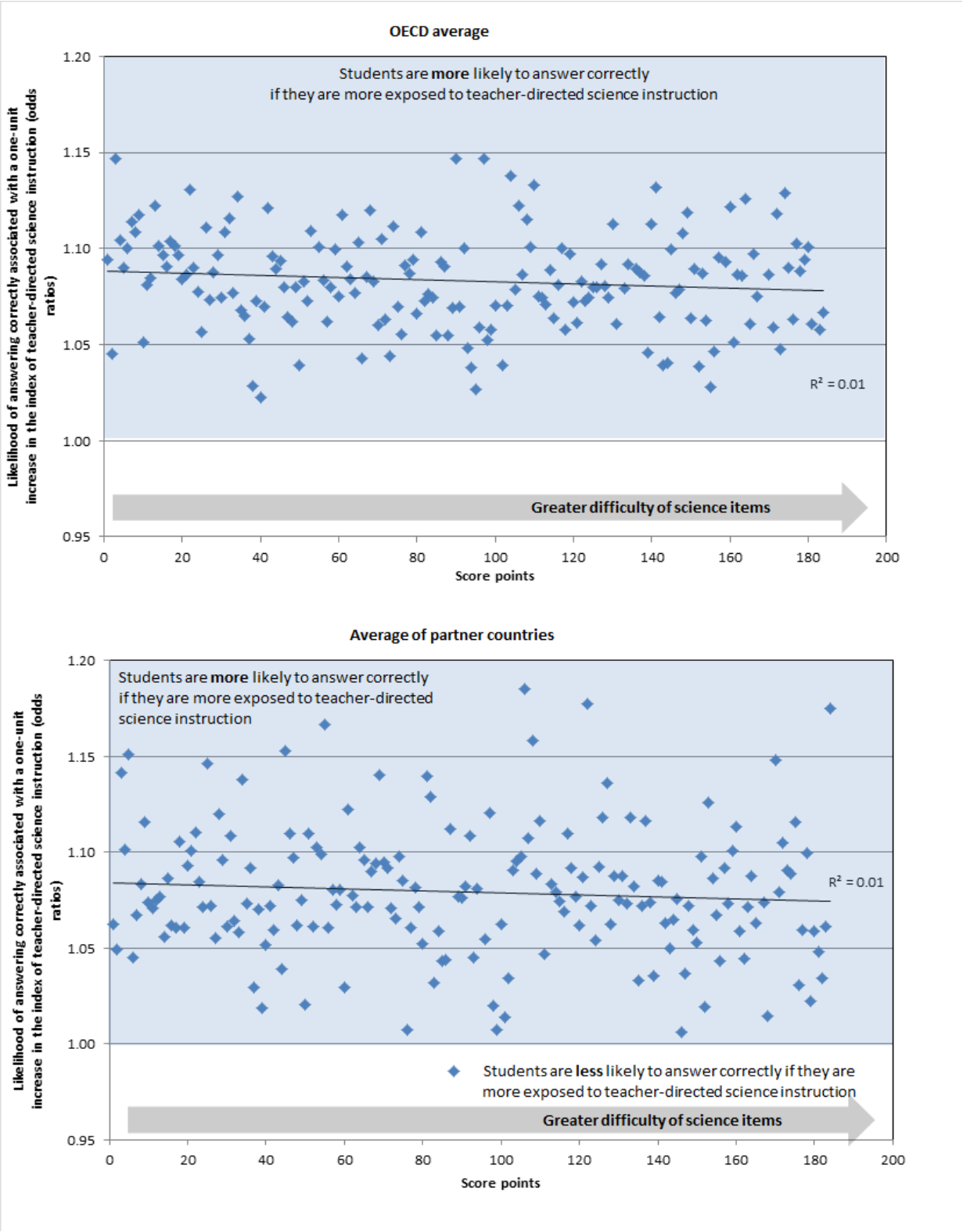
Put differently, combining the two teaching practices could be beneficial only if the interaction between EBST and TDSI is strongly and positively associated with science performance. The association should be strong enough to compensate for the negative effect of EBST. This is not the case here, as the regression coefficient on the interaction between EBST and TDSI is not significant in most countries, except Croatia, Czech

Republic, Estonia, Kosovo, Lithuania, Montenegro, Peru, Romania, the Slovak Republic, Thailand and Turkey, where it is positive, significant and weak (i.e. an improvement of between one and four score points in science performance).

Does TDSI help students engage with difficult science tasks?

Logistic regressions were carried out with the objective of determining whether exposure to TDSI helps students correctly answer science questions of varying difficulty. Figure 4.10 shows that while TDSI helps students answer science questions correctly, this association does not vary according to the level of difficulty of the question. A rise of one unit on the index of exposure to TDSI is associated with a rise of about 8% in the likelihood of answering a science question correctly regardless of the difficulty of the question. Similar findings were observed in OECD and in partner countries and economies.

Figure 4.10. Teacher-directed science instruction and student success on science items



Source: OECD, PISA 2015 Database.

Quantile regression analyses were used to determine whether the association between TDSI and science performance varies according to the level of student performance. The findings show no significant differences in this association among low, middle and high performers (students in the 10th, 50th and 90th performance percentiles, respectively). This finding holds true in all countries and economies.

Regression analyses investigating the interactions between TDSI and student and school characteristics also show no significant variations in the association between TDSI and science performance related to those variables. As in section 3, the analyses considered school disciplinary climate, students' experiences of teacher unfairness, sense of belonging at school, interest in and enjoyment of science, achievement motivation, shortages of school human and material resource, and teacher and student behaviour hindering learning.

The results of the three analyses show that the positive association between TDSI and science performance is not sensitive to students' level of performance, the difficulty of science tasks, or the school context. In other words, TDSI could be seen as an effective and consistent practice in transmitting scientific knowledge and skills; its success does not hinge on external factors beyond teachers' control.

5. Adaptive instruction in science lessons

In an episode of the popular television series *The Wire*, a teacher in a "distressed" school discovered the students playing poker games in the classroom. Instead of punishing them for gambling, the teacher grabbed the opportunity and used the poker games to teach statistics. The students learned without realising they were learning. Fortunately, most students are not as academically disengaged and emotionally distressed as those seen in that episode of *The Wire*, yet all teachers face some degree of student diversity in the classroom, and they all need to decide how much to adapt their lessons to the knowledge, skills and interests of students. This chapter focuses on how frequently science teachers around the globe are adapting the structure and content of their lessons, and how these practices are related to student outcomes.

Adaptive teaching is a pedagogical approach whereby teachers adapt the structure and content of the lesson to the characteristics of their students, such as their knowledge, skills and interests, and apply different teaching practices to different groups of learners (Borich, 2011). The goal is to engage all students academically and improve their understanding of the planned curriculum. According to Ikwumelu, Oyibe and Oketa (2015), adaptive teaching is different from overlapping teaching approaches, such as individualised and differentiated instruction, in that adaptive teaching is to a large extent a whole-class strategy. Adaptive teaching is also closely connected to student feedback since teachers usually revise their lessons following students' reactions and suggestions.

Teachers can adapt the lesson in several ways. When students are having trouble understanding a lesson, teachers can alter the curriculum, the structure of the lesson or the way the content is presented (Ikwumelu, Oyibe and Oketa, 2015). Often, these adaptations will also require adjusting the ways in which students' mastery of the content is assessed. Some examples of adaptive teaching practices include streamlining the curriculum (content), starting the lesson with games to engage students from the outset (structure) or introducing experiments to explain concepts that most students are failing to understand (presentation).

Compared to other teaching practices, the effectiveness of adaptive teaching has been evaluated less frequently. In a previous OECD report, it was already shown that, in most countries, students who reported that their science teachers adapt their teaching more frequently scored higher in science (OECD, 2016b). In stating that the most powerful feedback is the one that students provide to teachers, Hattie (2009) also hints at the benefits of adaptive teaching; after all, student feedback is used by teachers mainly to adapt their lessons and ensure learning is actually happening.

The risk of disregarding student feedback is that students lose interest and attention, and their sense of self-efficacy weakens. The problem of content overload is probably the best example: schools and teachers are forced to rush through a lengthy curriculum without ensuring that deep learning is actually taking place; many students may be lost, at least temporarily, along the way (Eylon and Linn, 1988; Schwartz, 2006). If the goal is to ensure that all type of students learn, adapting the lesson to students with different knowledge, skills and interests becomes an essential component of teaching (Hofstein and Lunetta, 2004).

In this chapter, PISA 2015 data is used to measure the frequency with which adaptive teaching practices are used in science lessons across school systems. The chapter also

examines how these practices relate to student outcomes, such as science performance, epistemic beliefs and students' expectations of pursuing a science-related career, and reveals the contexts where adaptive teaching practices are more likely to take place.

Box 5.1. Adaptive teaching in science lessons

The use of adaptive teaching was measured in PISA using student-reported information.

One question (question ST107) with three items was used. The question asked about the frequency of certain adaptive teaching practices undertaken in a specific science course that students had selected previously. Answers were provided on a four-point Likert scale ranging from "Never or almost never", "Some lessons", "Many lessons", to "Every lesson or almost every lesson". The question is:

> How often do these things happen in your lessons for this <school science> course?

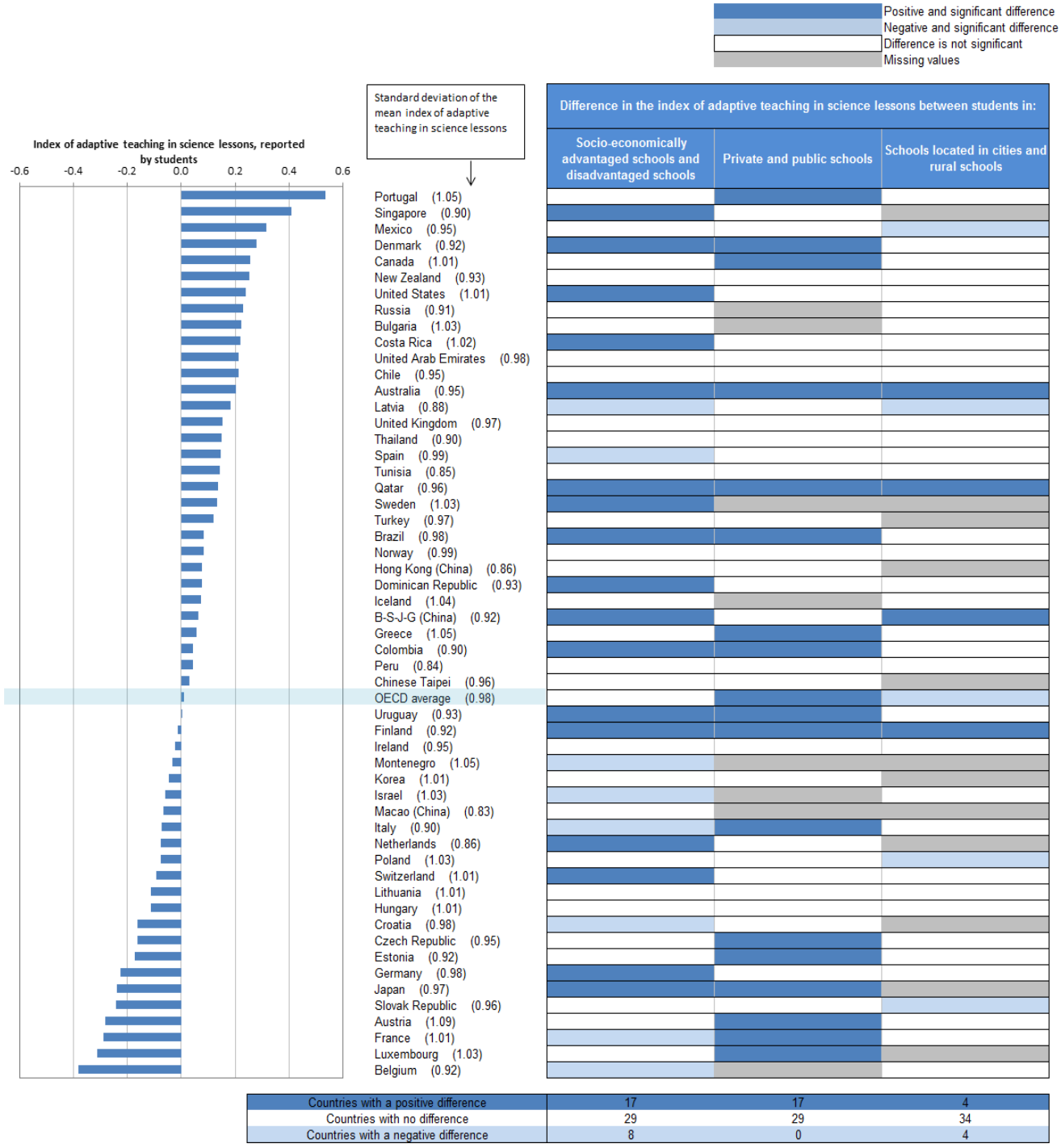
1. The teacher adapts the lesson to my class's needs and knowledge.
2. The teacher provides individual help when a student has difficulties understanding a topic or task.
3. The teacher changes the structure of the lesson on a topic that most students find difficult to understand.

An index was constructed based on the answers to these three statements using IRT scaling (see OECD, 2017c, *Technical Report*, Chapter 16). The index was standardised to have an average of 0 across OECD countries and a standard deviation of 1, meaning that two-thirds of the population fall between the values of -1 and 1 in the index.

5.1. Descriptive evidence on the use of adaptive teaching in science lessons

Figure 5.1 presents the average of the index of exposure to adaptive teaching in science lessons (hereafter "ATSL"). The findings show that, according to students' reports, science teachers in Canada, Denmark, Mexico, New Zealand, Portugal and Singapore adapt their lessons most frequently. In contrast, Austria, Belgium, France, Germany, Japan, Luxembourg and the Slovak Republic are among the countries where students are least exposed to adaptive teaching practices. In Belgium, for instance, about one in four students reported that their science teacher adapts the lesson to the needs and knowledge of the class in many lessons or every lesson, compared to six in ten students in Portugal who so reported (OECD, 2016b).

Figure 5.1. Index of adaptive teaching in science lessons, reported by students



	Socio-economically advantaged schools and disadvantaged schools	Private and public schools	Schools located in cities and rural schools
Countries with a positive difference	17	17	4
Countries with no difference	29	29	34
Countries with a negative difference	8	0	4

Note: Countries and economies are ranked in descending order of the index of adaptive teaching in science lessons, reported by students.
Source: OECD, PISA 2015 Database.

On average across OECD countries, students in socio-economically advantaged and disadvantaged schools reported similar exposure to ATSL (Figure 5.1). In 17 countries and economies, especially in Beijing-Shanghai-Jiangsu-Guangdong (China) (hereafter "B-S-J-G [China]"), Brazil, Japan, the Netherlands, Singapore and Sweden, teachers in advantaged schools adapt their teaching more frequently than teachers in disadvantaged schools. The reverse is observed in eight countries, including Croatia, France, Israel, Italy and Montenegro.

When comparing adaptive teaching practices by school location, the findings show that, on average across OECD countries, they are somewhat more prevalent in rural schools than in urban schools (Figure 5.1). In Latvia, Mexico, Poland and the Slovak Republic, science teachers in rural schools adapt their instruction more frequently, whereas in Australia, B-S-J-G (China), Finland and Qatar, science teachers in urban schools do, according to students' reports.

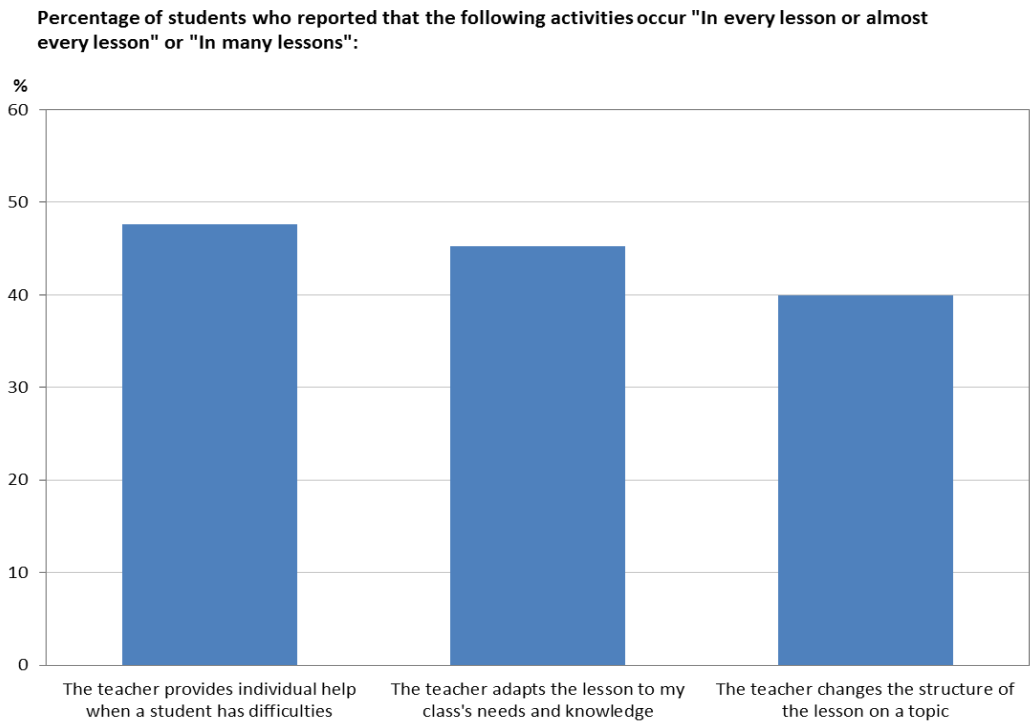
Students in private schools reported greater exposure to ATSL than students in public schools, on average across OECD countries and in 17 school systems. In Brazil, Denmark, Greece, Italy, Japan and Portugal, the difference between private and public schools on this index is more than 0.2 of a standard deviation.

PISA results show minimal differences in student-reported exposure to ATSL between schools where more than 30% of the student body have an immigrant background (first- and second-generation immigrant students combined) and schools where that proportion is smaller (Table 5.2). The same is true when considering schools where more than 30% of students speak a language at home that is different from the language of assessment and schools with smaller proportions of those students. This is surprising, given that, in many school systems, students with an immigrant background and those who speak a language different from the language of assessment show weaker academic performance (OECD, 2016b) and may require more flexible lesson plans.

According to students, their teachers use ATSL regularly. On average across OECD countries, almost half of students reported that their science teachers adapt the lesson to the needs and knowledge of the class, and provide support to struggling students in every, or many, lessons (Figure 5.2). Changing the structure of the lesson on a topic that students find difficult to understand is not as prevalent, but four in ten students reported that this happens in every, or many, lessons.

Figure 5.2. Adaptive teaching in science lessons

OECD average



Source: OECD, PISA 2015 Database.

5.2. Adaptive teaching in science lessons and students' attitudes and expectations

PISA 2015 examined students' enjoyment of science, interest in broad science topics, epistemic beliefs, and their expectation of having a science-related career when they are 30 years old. This sub-section explores the relationships between ATSL and students' attitudes and career expectations.

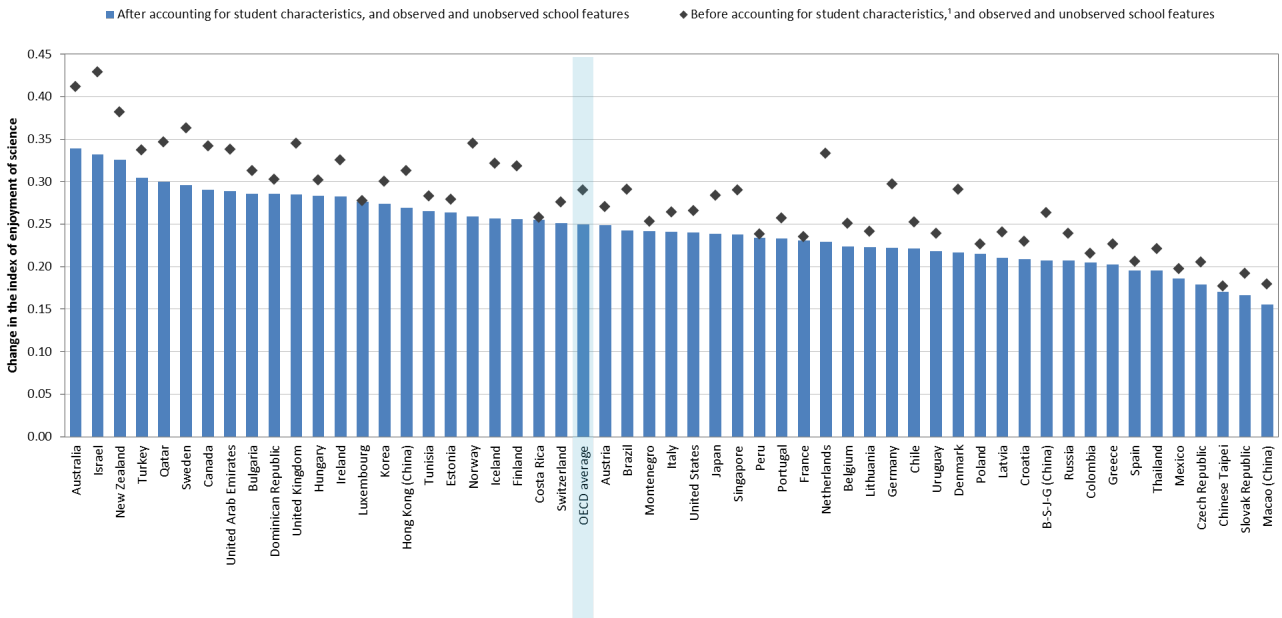
Enjoyment of science

Figure 5.3 presents the results of two regression analyses assessing the association between students' enjoyment of science and reported exposure to ATSL. The first analysis relies on simple regressions that do not account for any variable other than ATSL. The second relies on a school fixed-effects approach – which accounts for observed and unobserved school characteristics – that also accounts for students' socio-economic status, gender and grade level, and the number of science subjects taken. Students are expected to show greater engagement with a subject when teachers adapt the lesson plan to students' needs and choose tasks at an appropriate level of difficulty.

In every school system, students reported greater enjoyment of science when their science teachers adapted their lessons more frequently. The results are moderate but are stronger than for all other teaching practices, both before and after accounting for students' characteristics and school effects. The associations are particularly strong (regression coefficients above 0.3 in the school fixed-effects model) in Australia, Israel, New Zealand, Qatar and Turkey.

Figure 5.3. Enjoyment of science and adaptive teaching in science lessons

Change in the index of enjoyment of science associated with a one-unit increase in the index of adaptive teaching in science lessons, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: All values are statistically significant.
Countries and economies are ranked in descending order of the change in the index of enjoyment of science associated with a one-unit increase in the index of adaptive teaching in science lessons reported by students, after accounting for student characteristics, and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

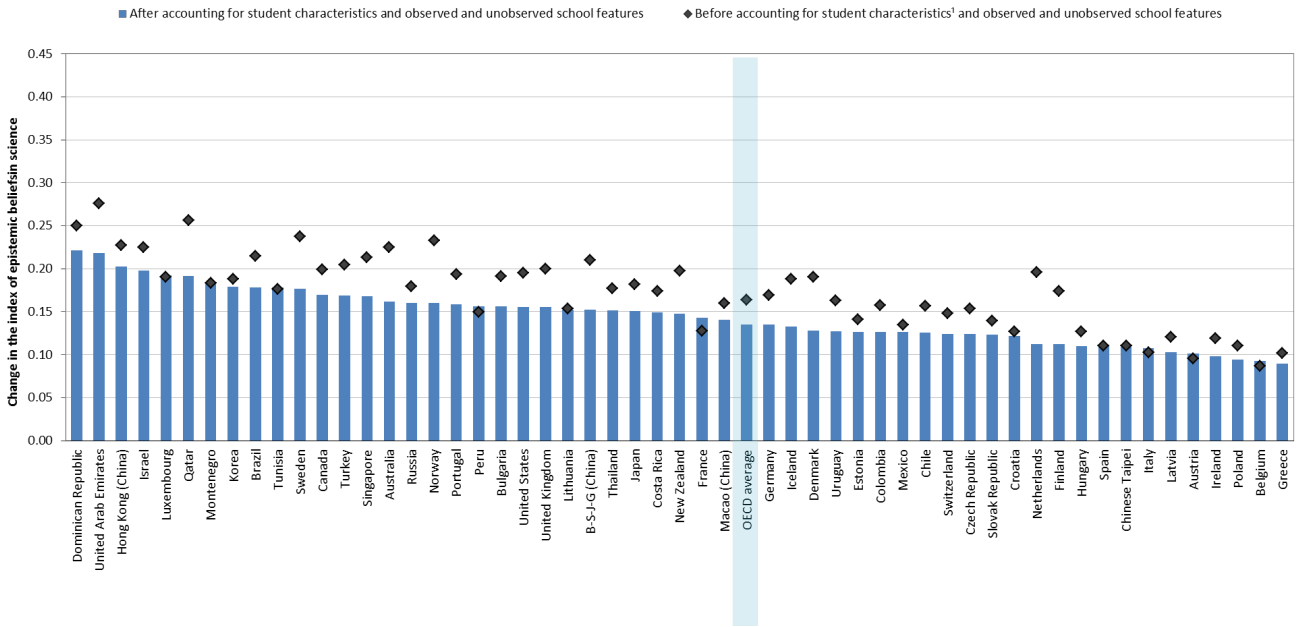
Similar positive but weaker associations (Table 5.4) are observed between interest in broad science topics and exposure to ATSL. The regression coefficients vary between 0.11 and 0.26 for all countries, after accounting for students' characteristics and school effects.

Science epistemic beliefs

This sub-section examines the relationship between students' epistemic beliefs and their self-reported exposure to ATSL. The results in Figure 5.4 show that, in every school system, students who reported greater exposure to ATSL hold stronger epistemic beliefs. Even though the associations with students' epistemic beliefs are weak to moderate (between 0.1 and 0.25 in most school systems, after accounting for students' characteristics and school effects), they are the strongest among all teaching practices. In all countries and economies, the associations are weak (below 0.2), except in the Dominican Republic, Hong Kong (China) and the United Arab Emirates, where they are moderate (from 0.2 to 0.3).

Figure 5.4. Epistemic beliefs in science and adaptive teaching in science lessons

Change in the index of epistemic beliefs in science associated with a one-unit increase in the index of adaptive teaching in science lessons, before and after accounting for student characteristics and observed and unobserved school features



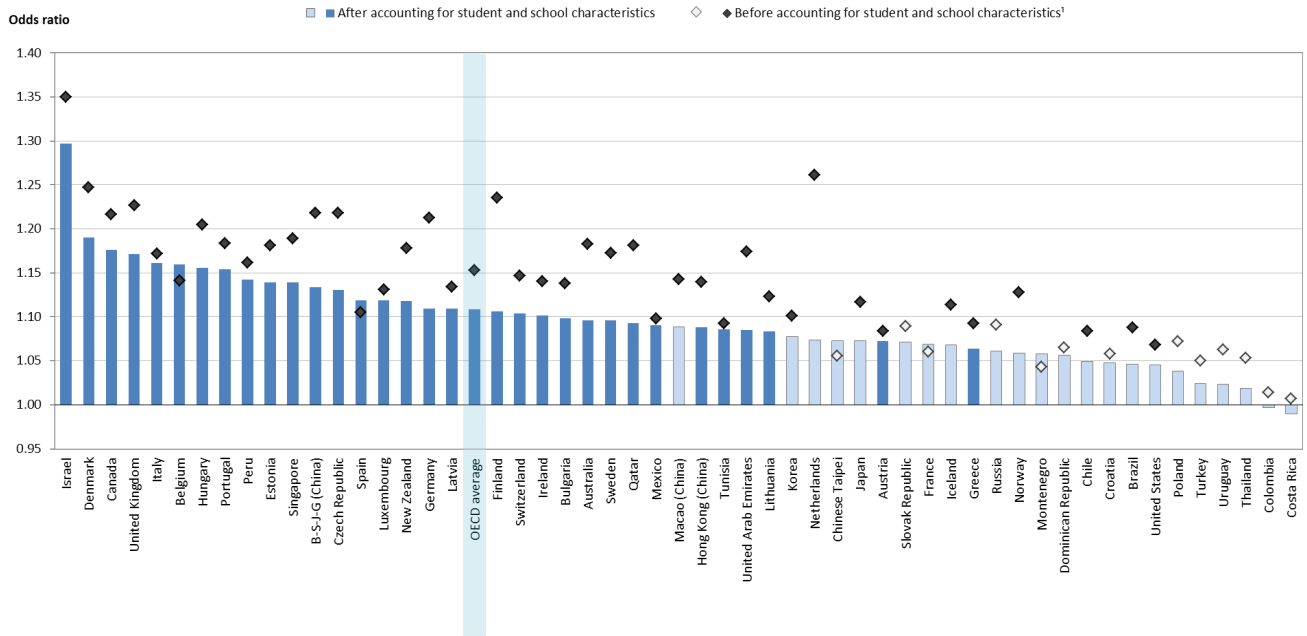
1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: All values are statistically significant.
Countries and economies are ranked in descending order of the change in the index of epistemic beliefs in science associated with a one-unit increase in the index of adaptive teaching in science lessons reported by students, after accounting for student characteristics, and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

Expectations of a science-related career at age 30

In a majority of school systems, students whose teachers adapt their lessons more frequently were more likely to see themselves working in a science-related job, after accounting for students' characteristics and school effects (Figure 5.5). On average across OECD countries, for instance, a one-unit increase in the index of ATSL is associated with a 11% increase in the odds ratio of expecting to pursue a science-related career. The associations are strongest in Belgium, Canada, Denmark, Hungary, Israel, Italy, Portugal and the United Kingdom (an increase of at least 15% in the odds ratio per unit increase in the index), and are not significant in 22 countries and economies.

Figure 5.5. Expectations of a science-related-career and adaptive teaching in science lessons

Likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase in the index of adaptive teaching in science lessons, before and after accounting for student and school characteristics



1. Student and school characteristics include gender, science performance, the number of science courses attended, the grade in which the student is enrolled, and socio-economic status (as measured by the PISA index of economic, social and cultural status) at the student and school levels.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase in the index of adaptive teaching in science lessons reported by students, after accounting for student and school characteristics.
Source: OECD, PISA 2015 Database.

The findings from this section show that greater exposure to ATSL is positively related to students' enjoyment of and interest in science, and fosters their epistemic beliefs and expectations of a science-related career.

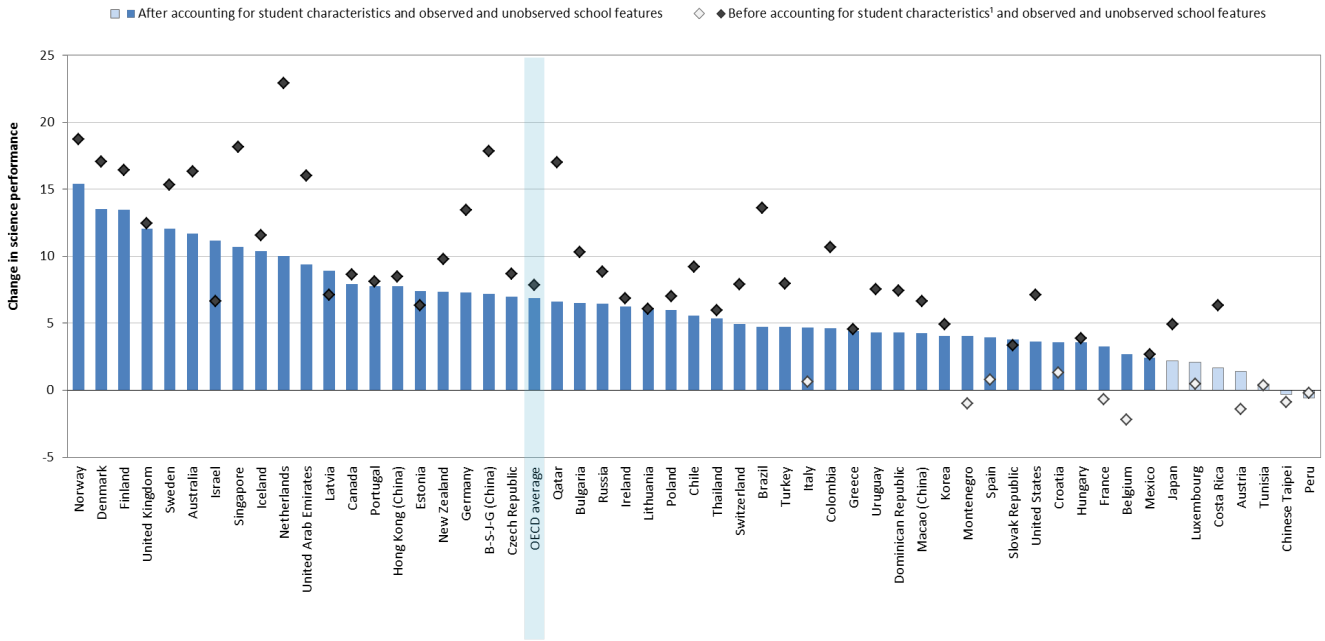
5.3. Adaptive teaching in science lessons and performance in science

This section examines the relationship between students' reports on the frequency of adaptive teaching practices in their science lessons and their science performance. Adapting lesson plans to the needs and knowledge of students, and providing individual help to struggling students may increase or limit students' learning opportunities, depending on the way in which these practices are carried out in practice.

In the majority of countries and economies, students who are more exposed to adaptive teaching practices in science lessons tend to score higher in science, both before and after accounting for students' characteristics and school effects (Figure 5.6). The association is particularly strong in the Nordic countries – Denmark, Finland, Iceland, Norway and Sweden – and in Australia, Israel, the Netherlands, Singapore and the United Kingdom. In Norway, for instance, a one-unit increase in the index is associated with a 15-point increase in the science score.

Figure 5.6. Performance in science and adaptive teaching in science lessons

Change in science performance associated with a one-unit increase in the index of adaptive teaching in science lessons, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, the number of science courses attended and the grade in which the student is enrolled.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase in the index of adaptive teaching in science lessons reported by students, after accounting for student characteristics, and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

The same analyses were repeated for all science subscales – those based on science competencies, knowledge categories and content areas – with similar results for all of them (the scales are discussed in detail in Annex A). On average across OECD countries, students always score seven points higher in science per one-unit increase in the index of ATSL, after accounting for students' characteristics and school effects.[7]

5.4. What motivates science teachers to adapt their teaching?

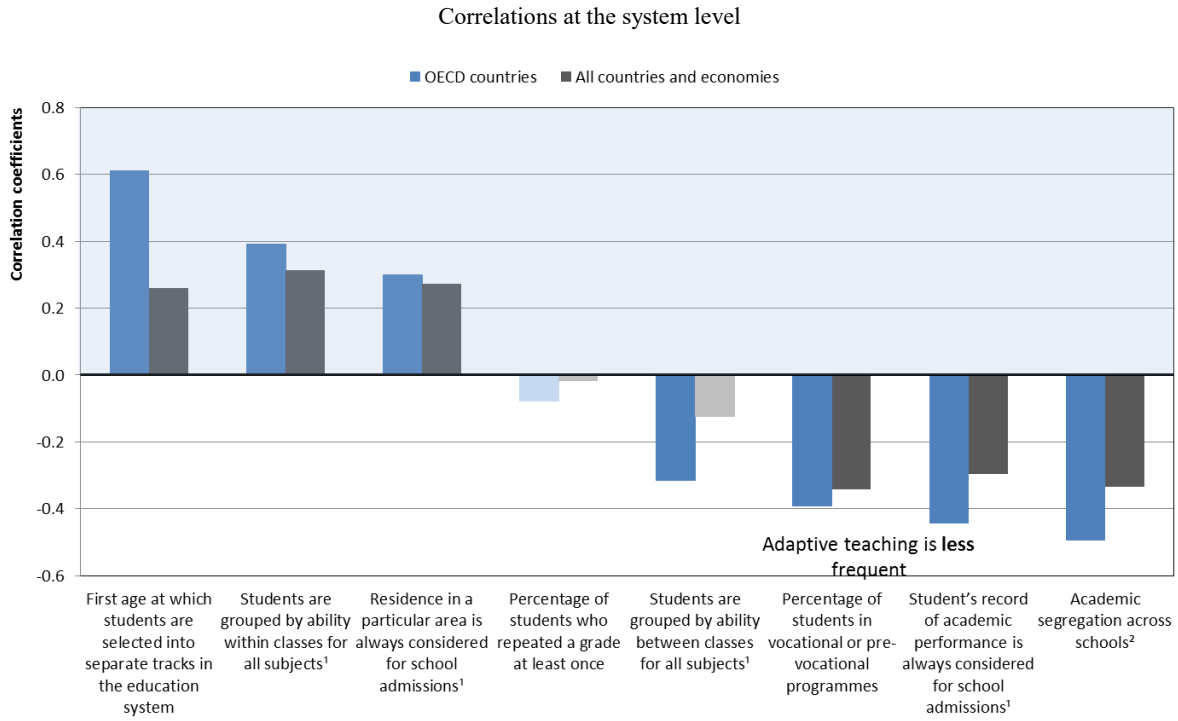
Given the positive associations between adaptive teaching and several student outcomes, such as science performance, epistemic beliefs, and the expectations of completing a university degree and pursuing a science-related career, it is essential to understand what enables or motivates teachers to adapt their lesson structure to students' needs and knowledge. For instance, as previously noted, teachers in private schools use adaptive teaching more frequently than do teachers in public schools. This section also analyses the following school factors: stratification, school size, the disciplinary climate in science lessons, and the diversity of the student body.

Stratification

Stratification refers to the various ways that education systems and schools organise instruction for students of varying ability, behaviour, interests and pace of learning (Dupriez, Dumay and Vause, 2008). In comprehensive systems, students of the same age follow a similar path until the end of compulsory education. In stratified systems, students of different abilities, behaviour or interests are separated into different grade levels, schools or classes. In stratified systems, classrooms tend to be, almost by design, more homogeneous and teachers can follow a more rigid lesson structure. By contrast, in comprehensive systems, there is greater diversity in the classroom, meaning that adapting lessons and supporting struggling students become essential elements of teaching.

In line with these expectations, Figure 5.7 shows that in school systems with greater stratification – whether this means selecting students into different tracks at a young age, admitting students to school based on their previous academic performance, segregating schools academically or grouping students by ability into separate classes – students were less likely to report that their science teachers adapted their lessons to students' needs and knowledge, and provided individual help to struggling students. By contrast, grouping students by ability within classes or admitting students into schools based on their place of residence is associated with greater frequency of adaptive teaching in science lessons.

Figure 5.7. Stratification and adaptive teaching in science lessons

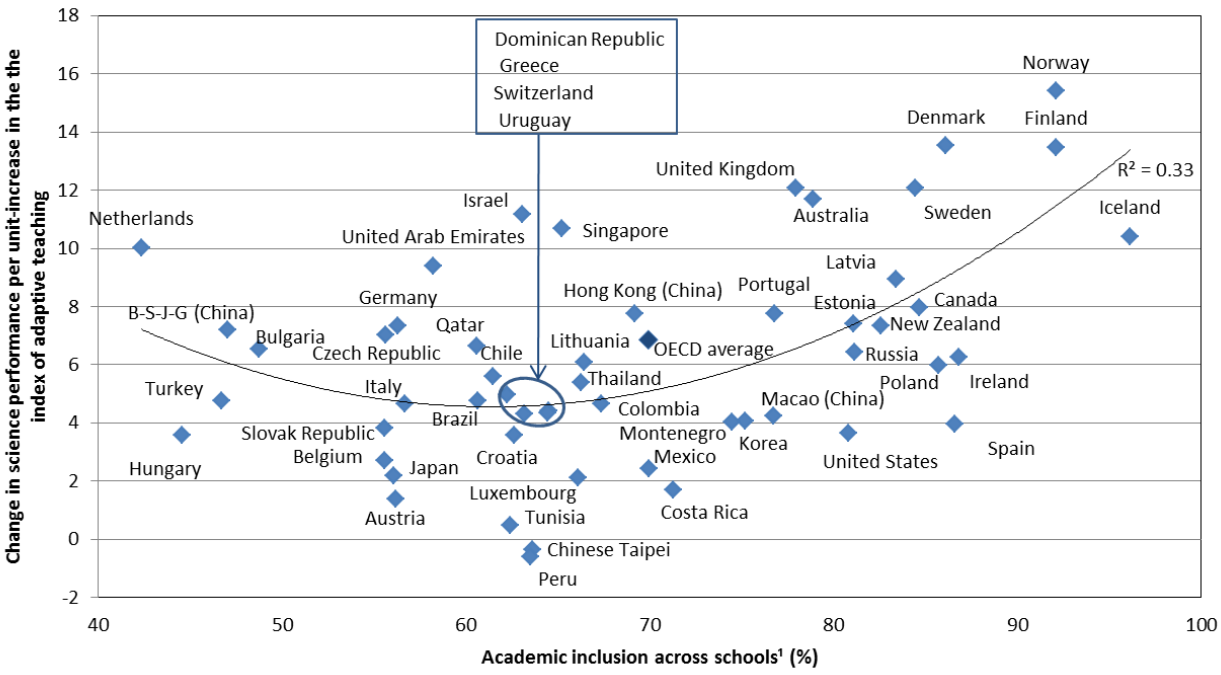


1. Percentage of students in schools where…
2. Academic segregation refers to the percentage of variation in science performance explained by *schools.*
Notes: Analysis based on 33 OECD countries, and 20 partner countries and economies.
Statistically significant correlations are marked in a darker tone.
Source: OECD, PISA 2015 Database.

The correlation is particularly strong regarding the age at which students are selected into different educational paths: the later students are separated into different tracks for the first time, the more likely they were to report that their science teachers adapted their lessons. Countries with the lowest average values in the index of adaptive teaching were those that track students at an early age (14 or younger), such as Austria, Belgium, Croatia, the Czech Republic, Germany, Hungary, Italy, Luxembourg, the Netherlands, the Slovak Republic and Switzerland (Figure 5.1). A previous OECD report reveals a similar relationship between early tracking and the index of teacher support (OECD, 2016b).

Not only is adaptive teaching more frequently used in comprehensive than in stratified school systems, the positive association with science performance is also stronger in countries and economies with greater academic inclusion (Figure 5.8). For instance, in Denmark, Finland, Iceland, Norway and Sweden – all highly inclusive school systems – students score at least 10 points higher in science per one-unit increase in the index of ATSL, compared to 7 points higher on average across OECD countries. However, the association is also strong in a few countries with relatively large between-school differences in science performance, such as Israel, the Netherlands and the United Arab Emirates.

Figure 5.8. Association between science performance and adaptive teaching by education systems' stratification



1. Academic inclusion is equal to 100 – percentage of variation in science performance explained by schools.
Source: OECD, PISA 2015 Database.

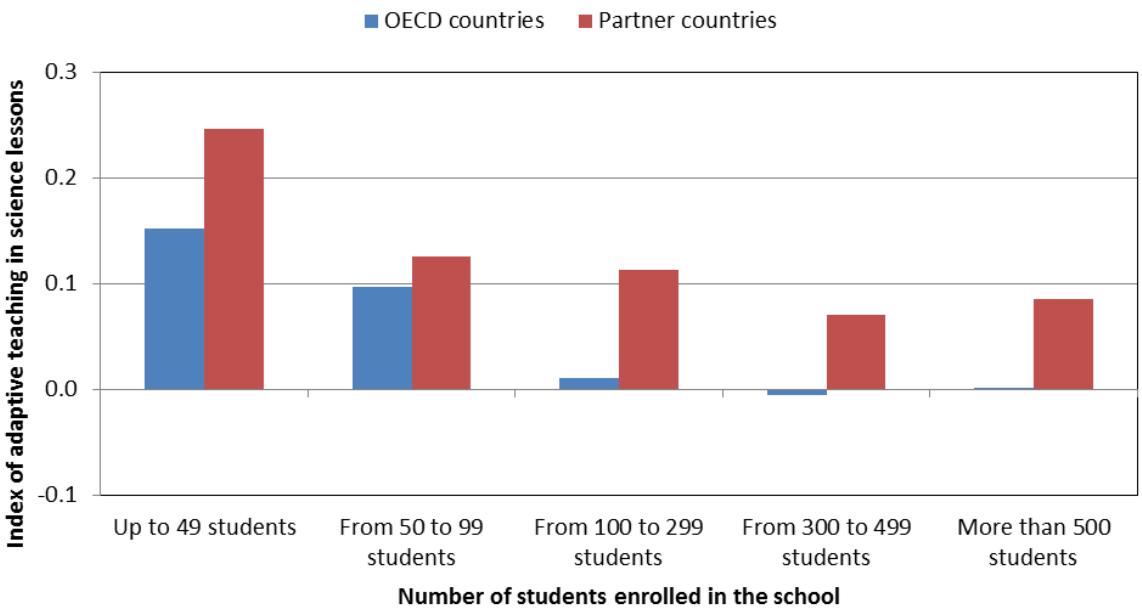
School size

Small schools may be financial burdens for education systems, but they also create plenty of opportunities in the area of teacher support and adaptive teaching, particularly for struggling and disadvantaged students (Duncombe and Yinger, 2007; Konstantopoulos and Chung, 2009; Schafft, 2016). Teachers in small schools are expected to adapt their lessons

and support their students to a greater extent than their counterparts in larger schools, among other reasons because they have greater opportunities to interact with all students and, especially in primary education, can do so over the course of several years.

PISA only looks into the learning experiences of 15-year-olds, but even in the secondary schools they attend, adaptive teaching is more frequently used in smaller schools than in larger schools. As Figure 5.9 shows, the index of adaptive teaching decreases with the size of the school, but only until a certain point – about 100 students across OECD countries, and 300 students across partner countries. Students are thus more likely to be exposed to adaptive teaching only in very small schools.

Figure 5.9. School size and adaptive teaching in science lessons



Source: OECD, PISA 2015 Database.

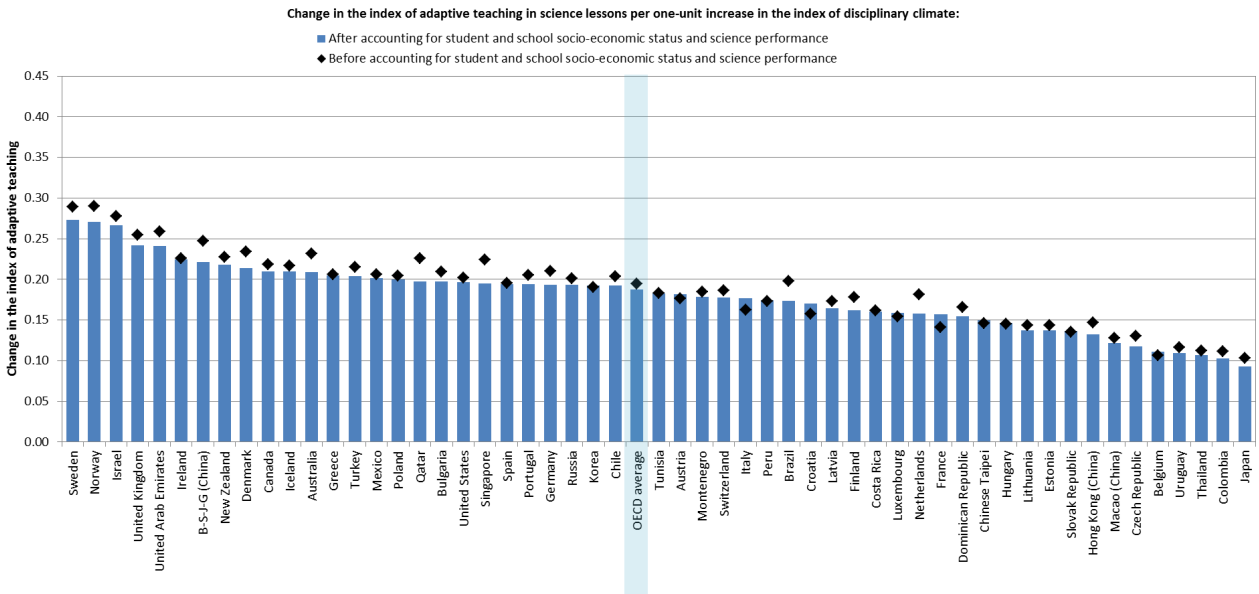
Disciplinary climate

An orderly classroom environment – one with limited noise, where students listen and behave – is essential for student learning (Willms and Ma, 2004). Only in such conditions can teachers successfully use diverse teaching strategies (see Chapter 3); only in such conditions can teachers adapt their teaching and support struggling students. PISA asked students how frequently the following things happen in their science 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 disciplinary climate whose average is zero and standard deviation is one across OECD countries.

Figure 5.10 shows that students who reported a more positive disciplinary climate in their science lessons were more likely to report that their science teachers adapt their teaching. This positive relationship is observed in every country and economy, before and after accounting for students' socio-economic status and science performance, and schools'

socio-economic profile. The strength of this relationship suggests that a positive disciplinary climate may be playing an enabling role for adaptive teaching practices.

Figure 5.10. Disciplinary climate and adaptive teaching in science lessons



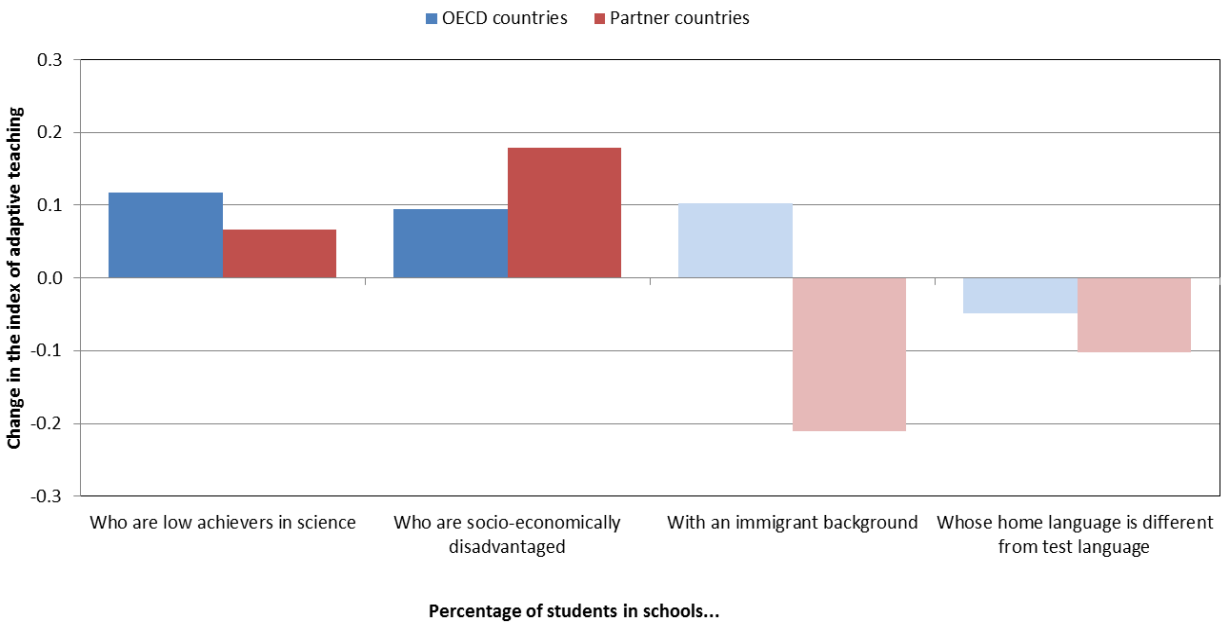
Note: All values are statistically significant.
Source: OECD, PISA 2015 Database.

Student diversity

When the student body is highly diverse teachers may have greater incentives to adapt their teaching. For instance, in classrooms with large proportions of low-achieving students, teachers may need to slow the pace of instruction, or even water down the curriculum, to ensure that all students engage with the activities and tasks proposed. Adaptive teaching may also be more frequently used in classrooms with higher numbers of socio-economically disadvantaged students and with large shares of students with an immigrant background or who speak a language different from the language of instruction.

PISA 2015 data show that students were more likely to report that their science teachers adapt their teaching when there are larger numbers of struggling students and disadvantaged students at the school, after accounting for students' gender, socio-economic status and science performance (Figure 5.11). However, as reported earlier, larger shares of students with an immigrant background at school, or of students who speak a language at home that is different from the test language, are not associated with changes in the index of adaptive teaching in science lessons. This does not necessarily mean that teachers do not respond to the needs of students with an immigrant background; it may simply mean that these students may only benefit from greater support and adaptive teaching when they are struggling academically and are socio-economically disadvantaged.

Figure 5.11. Student diversity at school and adaptive teaching in science lessons



Notes: All variables are included in a linear regression model accounting for students' gender, science performance and socio-economic status.
Statistically significant coefficients are marked in a darker tone.
Source: OECD, PISA 2015 Database.

Box 5.2. Teacher support

Student-teacher relationships in the classroom directly affect how students learn and how they feel towards school. When students feel supported by their teacher, not only do they perform at higher levels, they also feel more motivated and more satisfied about their lives (Guess and McCane-Bowling, 2016; Pitzer and Skinner, 2017).

Teacher support consists of providing students with an environment in which they feel that their teachers care about their learning and achievement, and that they receive the necessary help to succeed. Students who perceive that their teacher is emotionally supportive tend to be more engaged in school (Federici and Skaalvik, 2014). They pay more attention in class and prepare for lessons, which, in turn, leads to higher academic achievement (Klem and Connell, 2004).

Several studies also show that students who perceive their teacher as supportive of their learning are more motivated for academic activities (Ricard and Pelletier, 2016). Federici and Skaalvik (2014) find that teacher support in the form of additional help and explanations is associated with higher levels of intrinsic motivation.

In PISA 2015, teacher support in science lessons (TSSL) was assessed using student-reported information on one question (question ST100) with five related statements. The question asked about the frequency with which certain practices were undertaken by teachers in science classes: "The teacher shows interest in every student's learning"; "The teacher gives extra help when students need it"; "The teacher helps students with their learning"; "The teacher continues teaching until students understand"; and "The teacher gives an opportunity to express opinions". Responses were scored on a four-point Likert scale ranging from "Never or hardly ever", "Some lessons", "Most lessons", to "Every lesson". An index was constructed based on the responses to these five statements using IRT scaling (see OECD, 2017, Technical Report, Chapter 16).

On average across OECD countries, more than two in three students reported that each of the five practices was used in most or every science lesson. In each country, a minimum of 40% of students reported this level of frequency of TSSL (Table 5.9). Teacher support is most frequently provided in Albania, Costa Rica, the Dominican Republic, Kazakhstan, Kosovo, Mexico, Moldova and Portugal.

Students in socio-economically disadvantaged schools tend to be more exposed to teacher support than those in advantaged schools. In 17 countries and economies, students in private schools are more exposed to TSSL than those in public schools; but in 47 countries and economies, there is no difference in exposure between public and private schools. Students attending schools located in rural areas are slightly more exposed to TSSL than students attending schools in cities; the reverse is true in only three countries (Table 5.10).

In all countries, an increase in exposure to TSSL (teacher support in science lessons) is associated with positive attitudes towards science. Having a teacher who is attentive to students' need for help in science lessons significantly increases students' enjoyment of science, even after accounting for school fixed-effects and student characteristics, including science performance. The association is the strongest in Albania, Australia and New Zealand (where the regression coefficient is above 0.30) (Table 5.11).

In most countries and economies, TSSL also appears to be positively associated with the likelihood of students expecting to have a science-related career at age 30. In Israel, Kosovo and Lebanon, students are more than 20% more likely to expect a science-related career if they are exposed to supportive practices in science lessons, after accounting for student characteristics and school observed and unobserved features (Table 5.15). Moreover, in 25 out of 55 countries and economies with available and comparable data, students who perceive greater support from their teacher are more likely to expect to pursue their education and complete a university degree.

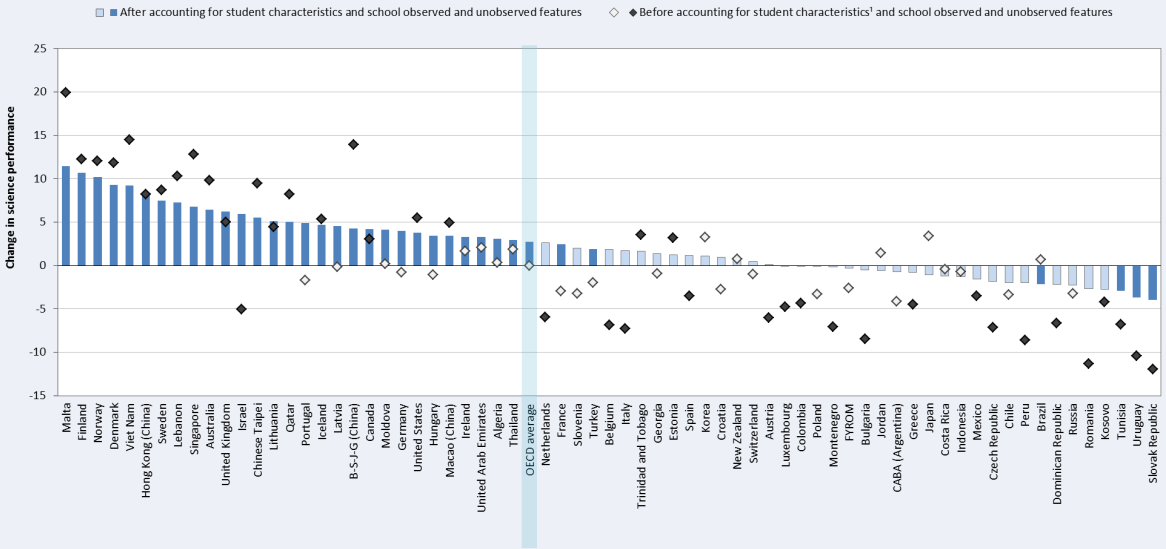
Teacher support and science performance are significantly associated in 41 countries and economies (Figure 5.12). However, in half of these, TSSL is negatively correlated with performance; in the other half, the association is positive. Once student characteristics, and observed and unobserved school features are taken into account, the relationship is no longer significant in almost all countries with previously negative coefficients. In other words, students who score lower are those who receive more teacher support, and their lower performance reflects other personal characteristics, such as their socio-economic status. Only in Brazil, the Slovak Republic, Tunisia and Uruguay is TSSL still negatively associated with performance after accounting for other factors, with students' scores two to four points lower when they are exposed to teacher support. In 31 countries and economies, students who are more exposed to TSSL score higher, after accounting for student characteristics, and observed and unobserved school features. They do so by at least 10 points in Finland, Malta and Norway.

PISA 2015 data show that boys and girls show different attitudes towards collaboration and teamwork (OECD, 2017b). Girls value relatively more the relational aspects of collaboration compared to boys, who value more the gains associated with teamwork. Based on these results, teacher support could be expected to have a stronger link to science performance among girls than among boys. However, despite the differences in attitudes, the association between TSSL and science performance does not vary according to gender, even after accounting for other characteristics, including performance.

The simultaneous exposure to teacher support and adaptive teaching practices is positively associated with science performance. Students in Mexico and Portugal are the most frequently exposed to both TSSL and ATSL. When considering the relationship between both practices combined and academic achievement (by adding an interaction between the two practices in the model), an increase of one unit in both indices is associated with an improvement in performance in almost all countries and economies. Students in Finland, Latvia and Norway who are more exposed to both practices combined score at least 15 points higher than their peers who are not as frequently exposed to these two practices. In these countries, and in more than 20 others, the association between ATSL and science performance is stronger when adaptive teaching practices are combined with teacher support, than when using ATSL alone (Table 5.17).

Figure 5.12. Performance in science and teacher support in science lessons

Change in science performance associated with a one-unit increase in the index of teacher support in science lessons, before and after accounting for student characteristics and school observed and unobserved features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, the number of science courses attended and the grade in which the student is enrolled.

Note: Statistically significant values are marked in a darker tone.

Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase in the index of teacher support in science lessons reported by students, after accounting for student characteristics and school observed and unobserved features.

Source: OECD, PISA 2015 Database.

6. Feedback in science classes

In a world where a growing number of occupations are at risk of being computerised (Frey and Osborne, 2017), teachers can still have an advantage over computers. One of these advantages may be providing meaningful and strategic feedback to students. Computers can certainly provide immediate feedback to students who engage in computer-assisted learning programmes, such as highlighting mistakes and indicating correct answers. However, teachers are the best placed to strategically adapt the content, timing and format of the feedback they provide to students because, among other reasons, they are usually informed about students' learning progress and personal context.

Teacher feedback is the information provided to students with the purpose of shaping their behaviour and closing the gap between actual and desired performance and understanding (Hattie and Timperley, 2007; Sadler, 1989; Tunstall and Gsipps, 1996). Examples of teacher feedback include grading students, marking assignments, telling students how to improve their performance and praising them for good behaviour.

Teacher feedback can be classified in multiple ways. For instance, feedback can be of a positive (e.g. praise and rewards), neutral (e.g. marks) or negative (e.g. criticism and punishment) nature, with some researchers even suggesting an optimal 3:1 ratio between positive and negative feedback (Voerman et al., 2012). Another common distinction is between academic feedback, usually related to the performance on a specific task, and behavioural feedback, of which boys typically receive the lion's share (Dweck et al., 1978; Jones and Dindia, 2004).

Researchers have also classified feedback into effective and ineffective practices, focusing on three main questions: "What are the goals?", "What progress is being made towards the goal?" and "What activities need to be undertaken to make better progress?" (Hattie, 2009; Hattie and Timperley, 2007; Kluger and DeNisi, 1996; Voerman et al., 2012). Hattie and Timperley (2007), for their part, consider that feedback practices operate at four levels: task, process, self-regulation and the self (i.e. personal evaluations and affect). Other common classifications are related to the format of feedback (oral, visual and written), its level of complexity, the type of recipient (individual or group), the reasons attributed to task success (effort, ability, chance and task difficulty), or the stage in the learning process (progress and discrepancy feedback).[8]

Feedback certainly has a powerful influence on student learning. According to Hattie's (2009) synthesis of 800 meta-analyses (a statistical technique for summarising the findings from independent studies), feedback is among the top ten influences on student achievement: its effect size is twice as large as the average effect of just attending school.

However, not all types of feedback are equally powerful (Kluger and DeNisi, 1996; Hattie, Gan and Brooks, 2017). Actually, previous research has demonstrated that some feedback practices are unrelated to student achievement, and others may even hinder student learning. For instance, according to Hattie and Timperley (2007), praise, extrinsic rewards and punishment are largely ineffective practices – particularly for interesting tasks – because they contain little information about how to perform a task. There is also evidence showing that students' learning may suffer when teacher feedback is mostly of a negative nature – especially among students with low self-efficacy (Hattie and Timperley, 2007) – because it can negatively affect their self-esteem and motivation, and increase the

frequency of defensive reactions (Reinke, Herman and Newcomer, 2016; Voerman et al., 2012).

What then are the most effective types of feedback? According to Hattie and Timperley (2007), the most effective forms of feedback are content-related and goal-oriented, and therefore rich in information about how students perform a task, and what they should do to improve. Previous research also highlights the importance of targeting feedback at the appropriate level to ensure that students relate the content of the feedback to previously acquired knowledge and skills (Hattie and Timperley, 2007; Kulhavy, 1977). Black and Wiliam's findings (1998) report that written comments are more effective than marks. Other debates remain open. For instance, on the timing of feedback, some researchers point to the benefits of delayed feedback (Butler, Karpicke and Roediger, 2007) while others suggest that immediate feedback is better (Opitz, Ferdinand and Mecklinger, 2011). In an attempt to settle the debate, Hattie and Timperley (2007) conclude that delayed feedback is better when it is task-related, whereas immediate feedback is desirable when it is process-related (i.e. "the processes of understanding how to do a task").

The effectiveness of diverse types of feedback may also differ between education systems. Sully de Luque and Sommer (2000), for instance, argue that "collectivist" cultures, such as those found in East Asian countries, feel at ease with indirect and group-focused feedback, and individualist cultures, such as that in United States, prefer more direct and personalised feedback.

In this chapter, PISA 2015 data is used to measure the frequency with which teacher feedback is used in science lessons across school systems. The chapter also examines how these practices relate to student outcomes, such as science performance, epistemic beliefs and students' expectations of pursuing a science-related career. The last section tries to bring out the nuances in the negative association between the index of teacher feedback in science lessons (hereafter "TFSL") (Box 6.1) and performance in the PISA science assessment.

Box 6.1. Teacher feedback in science lessons

The extent to which science teachers provide feedback to students was measured in PISA using student-reported information.

One question (question ST104) with five items was used. The question asked about the frequency of certain teacher feedback practices undertaken in a specific science course that students had selected previously. Answers were provided on a four-point Likert scale, ranging from "Never or almost never", "Some lessons", "Many lessons", to "Every lesson or almost every lesson". The question was:

> How often do these things happen in your lessons for this <school science> course?

4. The teacher tells me how I am performing in this course.
5. The teacher gives me feedback on my strengths in the school science subject.
6. The teacher tells me in which areas I can still improve.
7. The teacher tells me how I can improve my performance.
8. The teacher advises me on how to reach my learning goals.

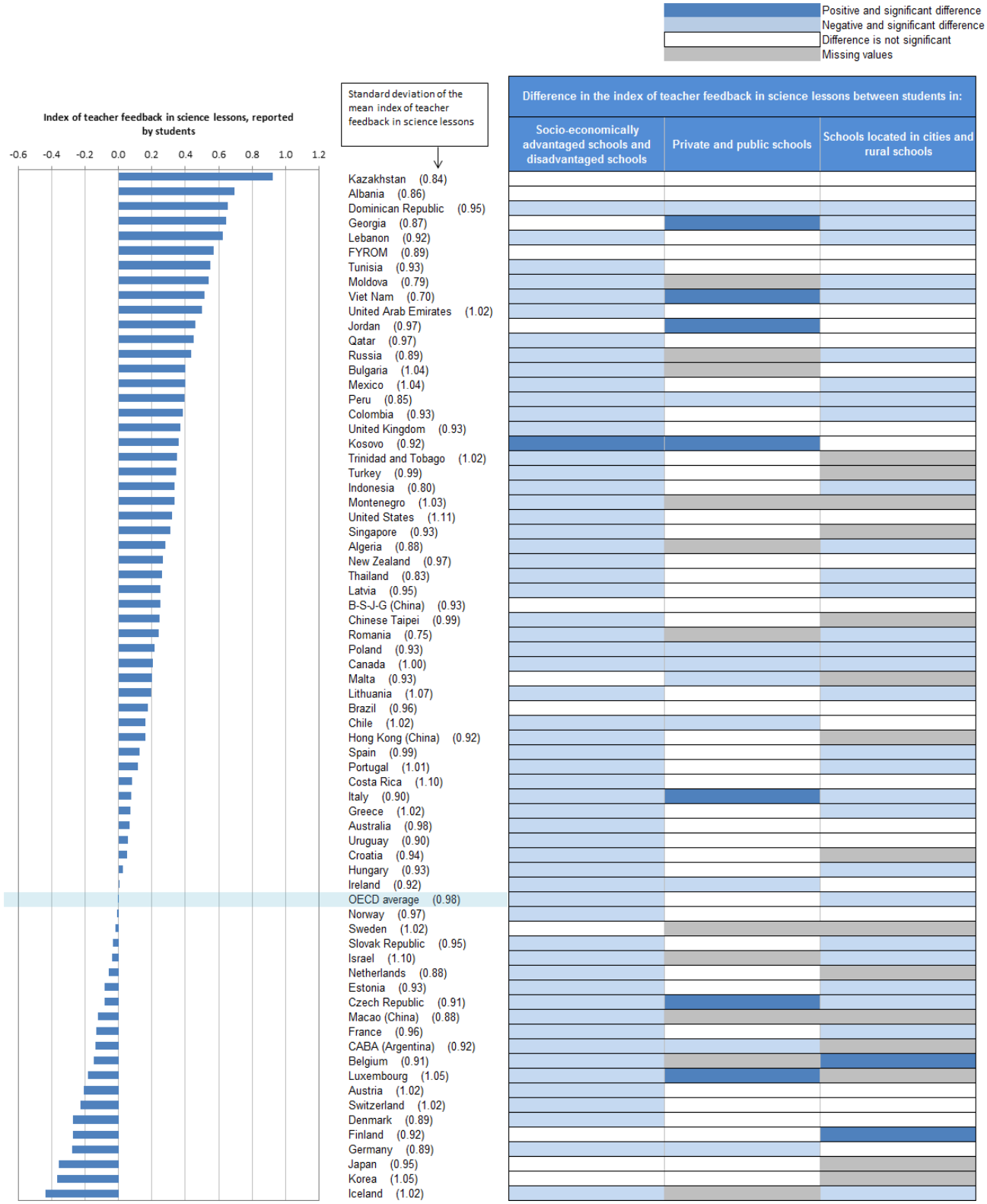
An index was constructed based on the answers to these five statements using IRT scaling (see OECD, 2017c, Technical Report, Chapter 16). The index was standardised to have an average of 0 across OECD countries and a standard deviation of 1, meaning that two-thirds of the population fall between the values of -1 and 1 on the index.

6.1. Descriptive evidence on the use of teacher feedback in science lessons

Figure 6.1 presents the average of the index of exposure to teacher feedback in science lessons. The findings show that, according to students' experiences, science teachers in Albania, the Dominican Republic, the Former Yugoslav Republic of Macedonia (hereafter "FYROM"), Georgia, Kazakhstan, Lebanon, Moldova, Tunisia, the United Arab Emirates and Viet Nam provide feedback most frequently. All of these countries have an average on the index which is at least 0.5 of a standard deviation higher than the OECD average. In contrast, students in Austria, Denmark, Finland, Germany, Iceland, Japan, Korea and Switzerland reported receiving the least frequent feedback. In Japan, for instance, about 1 in 6 students reported that their science teacher tells them how they are performing in the course in many lessons or every lesson, compared to one in two students in Georgia (OECD, 2016b).

On average across OECD countries and in 56 countries and economies, students in socio-economically disadvantaged schools reported greater exposure to TFSL than students in advantaged schools (Figure 6.1), whereas the reverse was observed only in Kosovo. In 12 school systems, including Finland, Japan and Korea, there was no statistically significant difference between advantaged and disadvantaged schools.

Figure 6.1. Index of teacher feedback in science lessons, reported by students



Source: OECD, PISA 2015 Database.

When comparing teacher feedback by school location, the findings show that, on average across OECD countries and in 28 education systems, feedback is more prevalent in rural schools than in urban schools (Figure 6.1). The rural-urban gap, in favour of rural schools, is particularly large in Hungary, Mexico and the Slovak Republic, all of which show considerable disparities in science performance between students in rural and urban schools (Echazarra and Radinger, forthcoming).

Differences in exposure to teacher feedback are small between public and private schools, according to students' reports. In only seven countries is teacher feedback more frequent in private schools than in public schools, and in nine countries and economies feedback is more frequent in public schools.

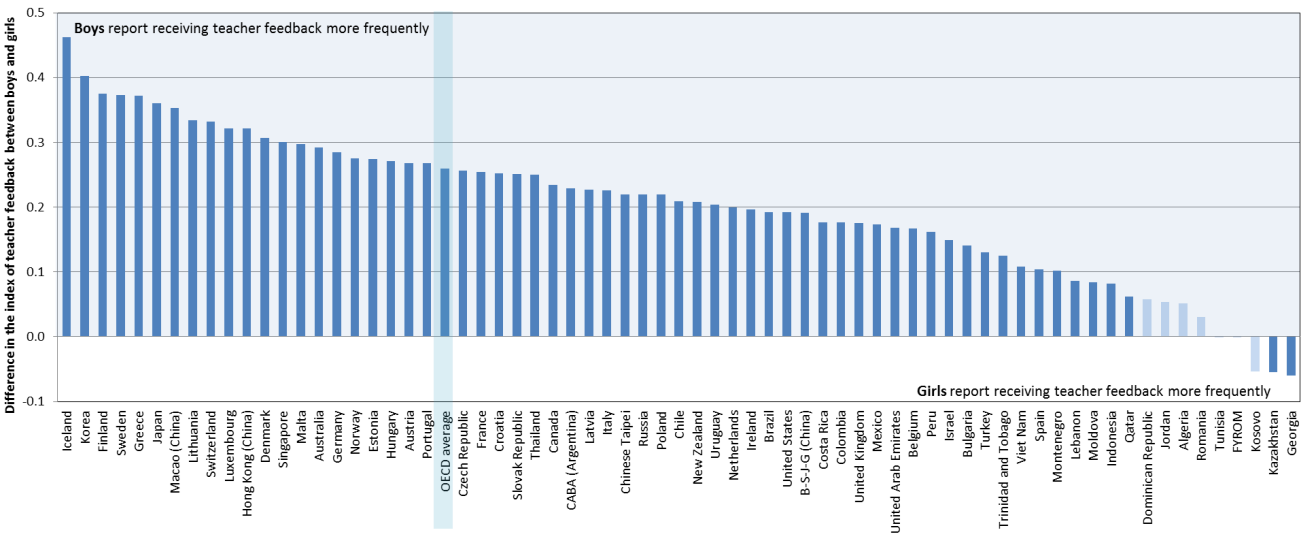
Teacher feedback also seems to be related to the composition of a school's student body. The results show greater exposure to TFSL in schools where more than 30% of students have an immigrant background (first- and second-generation immigrants combined) or who speak a language other than the language of assessment than in schools where less than 30% of students have an immigrant background (Table WP2_perfeed_sch).

Overall these results suggest that science teachers are providing more frequent feedback to students who need it the most: those attending disadvantaged and rural schools and those enrolled in schools with larger shares of students with an immigrant background.

Previous studies have generally found that boys receive more frequent teacher feedback, including praise, instruction and criticism, than girls (Dunbar and O'Sullivan, 1986; Kelly, 1988). This gender gap in teacher feedback led to an ongoing debate around the preferential treatment of boys by their teachers (Beaman, Wheldall and Kemp, 2006). However, while teachers' bias may be sometimes responsible for the gender imbalance in student-teacher interactions, according to some researchers (French and French, 1984) it is often the attention-seeking strategies and disruptive behaviour of only a few boys that are to blame. In addition, it is unclear how much these boys benefit from receiving constant feedback from their teachers given that it is largely negative.

In almost every country, boys reported receiving more frequent feedback than girls, with the largest disparities observed in Finland, Greece, Iceland, Korea, Japan and Sweden (Figure 6.2). Only in Georgia and Kazakhstan did girls report receiving more frequent feedback than boys. In addition, this gender gap barely changes after accounting for students' science performance, truancy, achievement motivation or sense of belonging at school.[9]

Figure 6.2. Gender differences in the frequency of science teachers' feedback

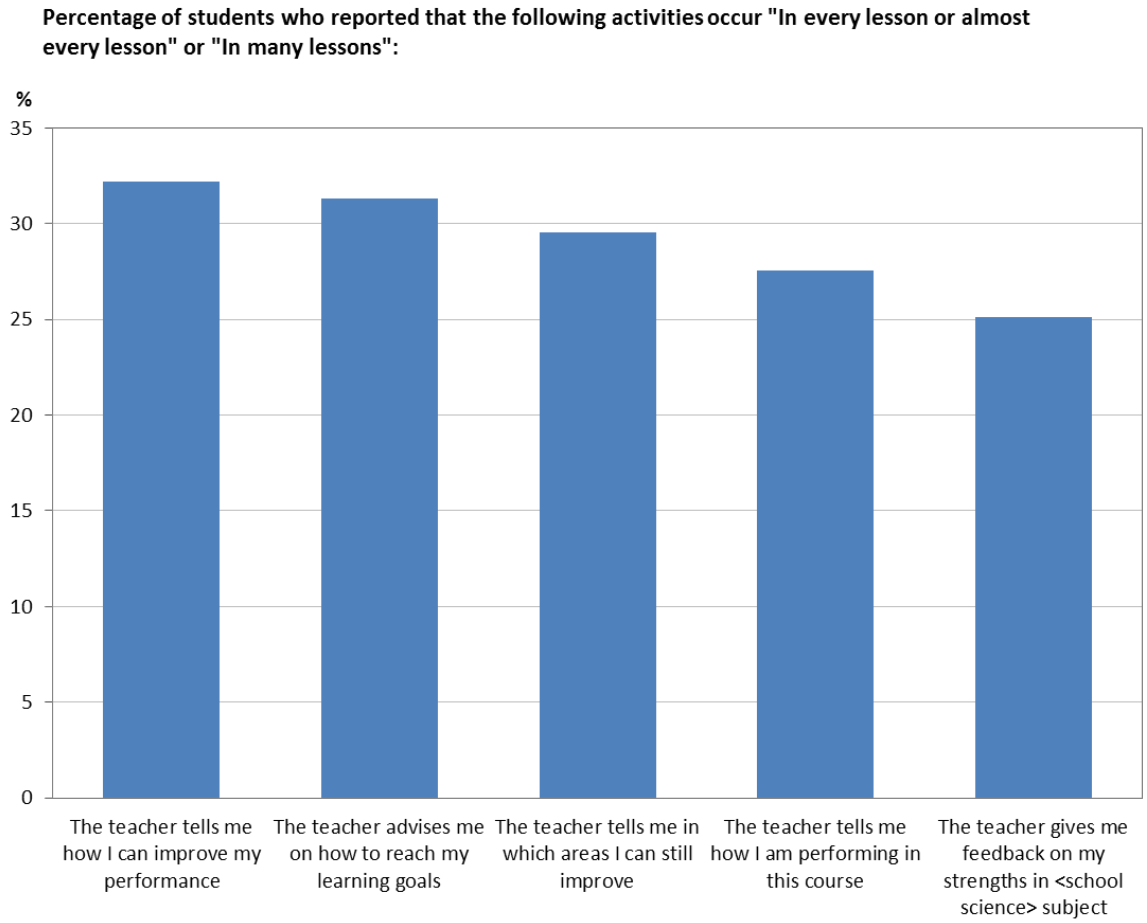


Note: Statistically significant differences are shown in a darker tone.
Countries and economies are ranked in descending order of the difference in the index of teacher feedback in science lessons between boys and girls.
Source: OECD, PISA 2015 Database.

The five types of feedback examined in PISA 2015 are given only sporadically, on average across OECD countries (Figure 6.3). This low incidence of feedback is consistent with previous research (Hattie and Timperley, 2007; Voerman et al., 2012). For instance, according to students' reports, about one in four students across OECD countries reported that their science teacher gives them feedback on their strengths in every, or in many, lessons. Perhaps surprisingly, on average across OECD countries, as much as 27% of students say their science teacher never, or almost never, tells them how they are performing in the course, and 32% of students reported that their science teacher never, or almost never, advises them on how they can reach their learning goals (OECD, 2016b). In Japan, as many as one in two students reported that their science teacher never, or almost never, tells them how they are performing in the course. The large share of students who reported that they do not receive any explicit feedback on their performance could be explained in several ways. For instance, these students may be good at self-regulating their learning, large classes may leave little time for teachers to provide individual feedback (Box 6.2) or, in certain education systems, secondary school teachers are simply not expected to provide individual feedback to students.

Figure 6.3. Teacher feedback in science lessons

OECD average



Source: OECD, PISA 2015 Database.

6.2. Teacher feedback in science lessons and students' attitudes and expectations

PISA 2015 examined students' enjoyment of science, interest in broad science topics, epistemic beliefs, and their expectation of having a science-related career when they are 30 years old. This sub-section explores the relationships between TFSL and students' attitudes and career expectations.

Enjoyment of science

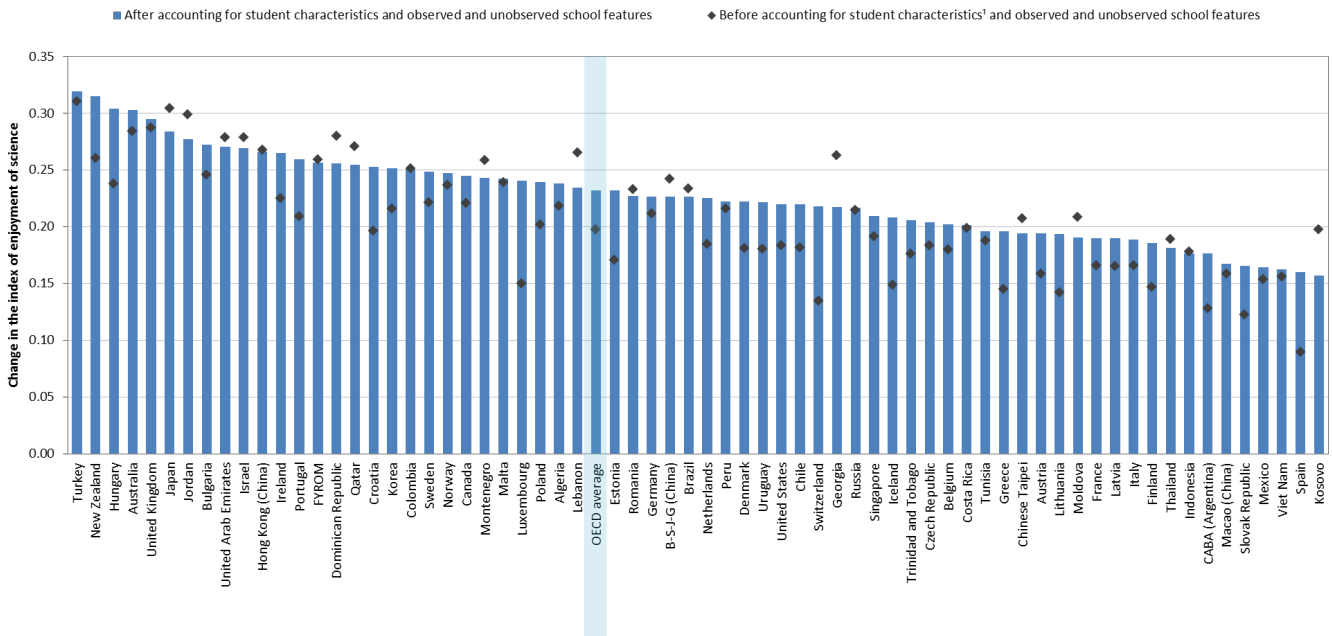
Figure 6.4 presents the results of two regression analyses assessing the association between student enjoyment of science and reported exposure to TFSL. The first analysis relies on simple regressions that do not account for any variable other than TFSL; the second relies on a school fixed-effects approach that also accounts for students' socio-economic status, gender and grade, and the number of science subjects taken. Since students are generally positive about receiving feedback from their teachers (Lee, 2008; Zhang, 1995), those who reported greater exposure to TFSL are expected to show greater enjoyment of science.

In every school system, students reported a greater enjoyment of science when their science teachers provide more regular feedback. The associations are moderate in magnitude and generally weaker than for other teaching practices, but they become stronger after accounting for students' characteristics and school effects. The associations are the strongest in Australia, Hungary, New Zealand and Turkey, and weakest in Kosovo, Mexico, Spain and Viet Nam.

Similar positive associations are observed between interest in broad science topics and exposure to TFSL. The regression coefficients vary between 0.12 and 0.25 for all countries and economies, after accounting for school effects and students' characteristics (Table WP2_intbrsci_perfeed).

Figure 6.4. Enjoyment of science and teacher feedback in science lessons

Change in the index of enjoyment of science associated with a one-unit increase on the index of teacher feedback in science lessons, before and after accounting for student characteristics and school observed and unobserved features



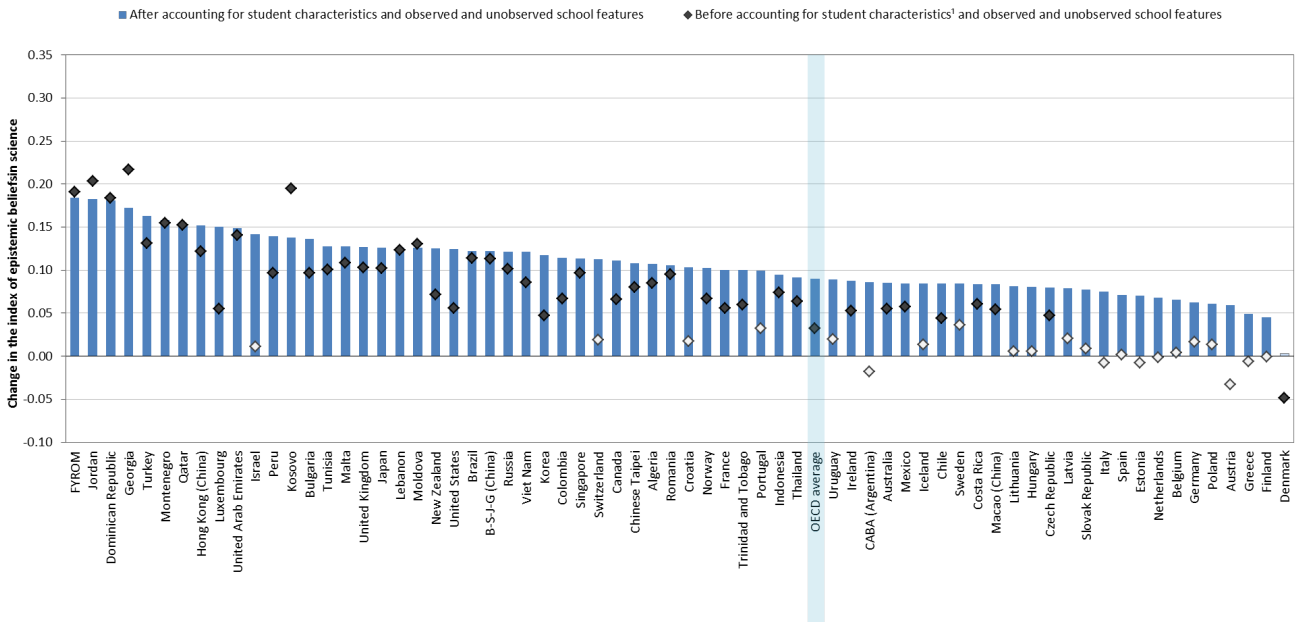
1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: All values are statistically significant.
Countries and economies are ranked in descending order of the change in the index of enjoyment of science associated with a one-unit increase on the index of teacher feedback in science lessons reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

Science epistemic beliefs

This sub-section examines the relationship between students' epistemic beliefs and their self-reported exposure to TFSL. The results in Figure 6.5 show that, on average across OECD countries, students who reported greater exposure to TFSL hold stronger epistemic beliefs. The associations are generally weak and not significant in several countries, but become significant in all countries, except Denmark, after accounting for students' characteristics and school effects.

Figure 6.5. Epistemic beliefs in science and teacher feedback in science lessons

Change in the index of epistemic beliefs in science associated with a one-unit increase on the index of teacher feedback in science lessons, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, science performance, the number of science courses attended and the grade in which the student is enrolled.
Note: Statistically significant values are marked in a darker tone..
Source: OECD, PISA 2015 Database.

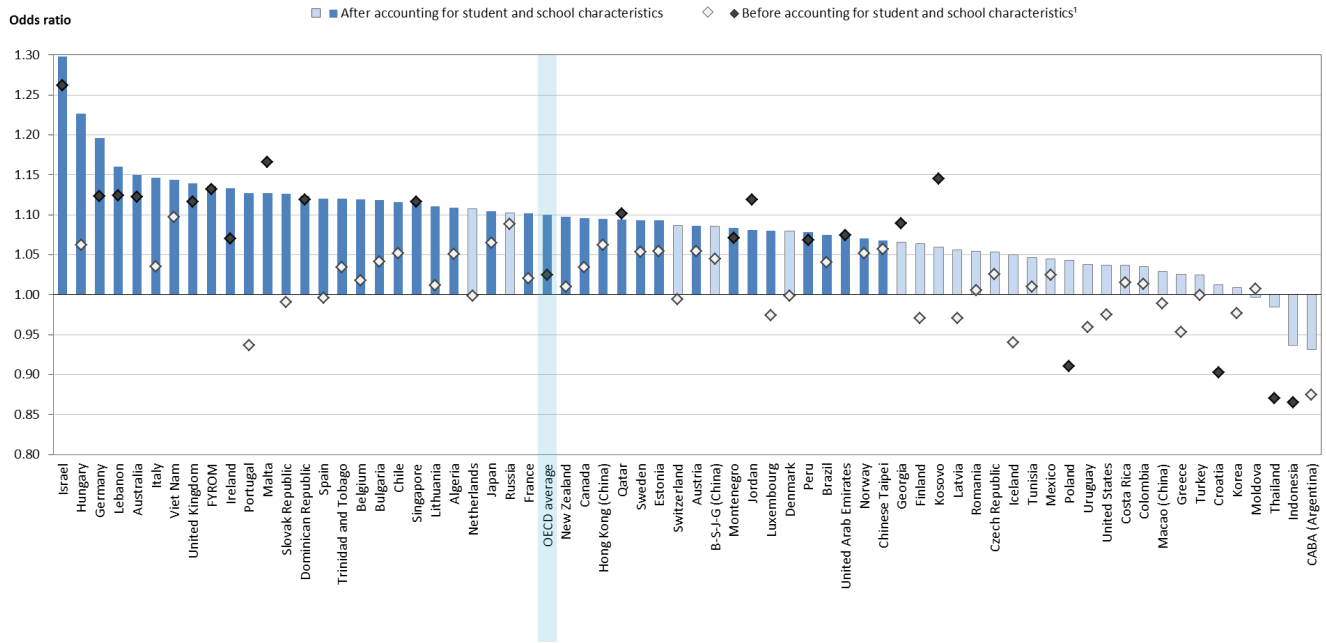
Expectations of a science-related career at age 30

In more than half of school systems, students whose science teachers provide feedback more frequently were more likely to see themselves working in a science-related job, after accounting for students' characteristics and school effects (Figure 6.6). On average across OECD countries, for instance, a one-unit increase in the index of TFSL is associated with a 10% increase in the odds ratio of expecting to pursue a science-related career. The associations are strongest in Australia, Germany, Hungary, Israel and Lebanon (an increase of at least 15% in the odds ratios per unit increase in the index), but are not significant in 28 countries and economies.

The findings from this section show that, on average across OECD countries, greater exposure to TFSL is positively related to students' enjoyment of and interest in science, and fosters their epistemic beliefs and expectations of a science-related career, particularly after accounting for students' characteristics and school effects.

Figure 6.6. Expectations of a science-related career and teacher feedback in science lessons

Likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase on the index of teacher feedback in science lessons, before and after accounting for student and school characteristics



1. Student and school characteristics include gender, science performance, the number of science courses attended, the grade in which the student is enrolled, and socio-economic status (as measured by the PISA index of economic, social and cultural status) at the student and school levels.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the likelihood of students expecting to work in a science-related career at age 30 associated with a one-unit increase on the index of teacher feedback in science lessons reported by students, after accounting for student and school characteristics.
Source: OECD, PISA 2015 Database.

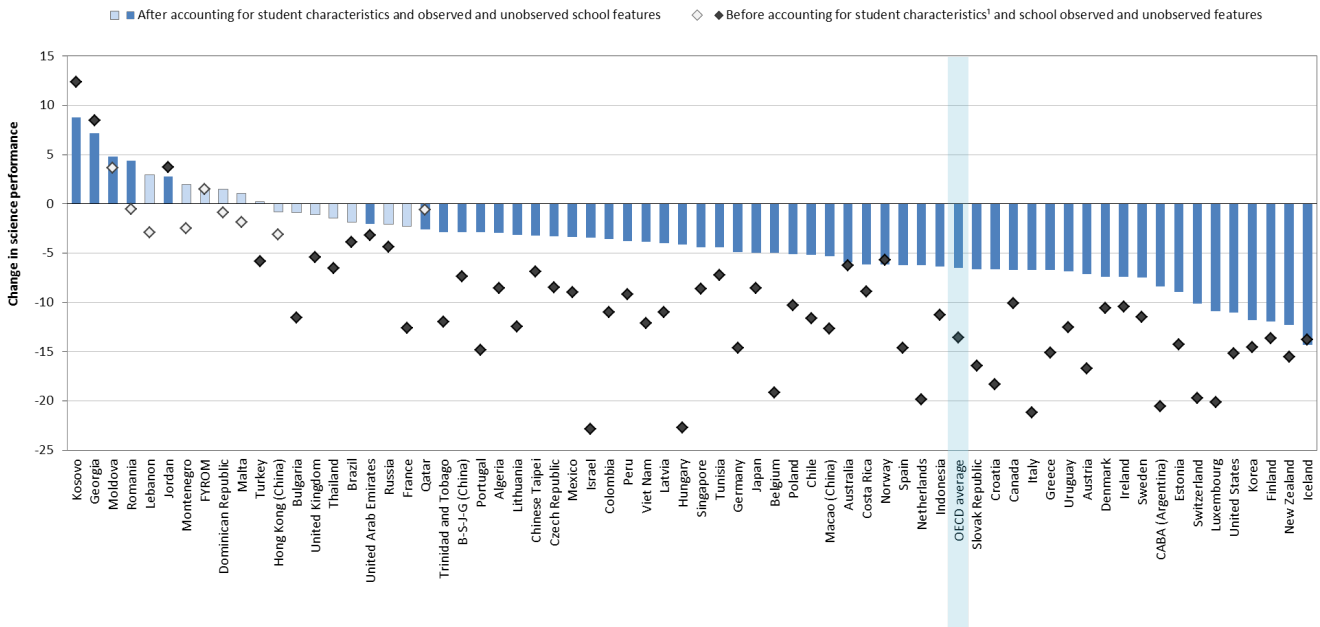
6.3. Teacher feedback in science lessons and performance in science

This section examines the relationship between students' reports on the frequency of teacher feedback in their science lessons and their science performance. As discussed in the introduction to this chapter, several studies have reported powerful effects of teacher feedback on learning outcomes. According to meta-analyses of a large number of education interventions, for instance, feedback is among the ten top influences on student achievement (Hattie, 2009). However, since teachers usually provide more frequent feedback to struggling students (Hyland, 2003), a cross-sectional study like PISA that collects data at one point in time may not produce the same results as randomised control trials, longitudinal analyses and other causal analyses.

The cross-sectional nature of PISA may explain why, in almost every country and economy, students who reported receiving teacher feedback more frequently tend to score lower in science, both before and after accounting for students' characteristics and school effects (Figure 6.7). For instance, in Finland, Iceland, Korea, Luxembourg, New Zealand, Switzerland and the United States, a one-unit increase in the index of TFSL is associated with a decrease of at least 10 score points in the science assessment.

Figure 6.7. Performance in science and teacher feedback in science lessons

Change in science performance associated with a one-unit increase on the index of teacher feedback in science lessons, before and after accounting for student characteristics and observed and unobserved school features



1. Student characteristics include socio-economic status (as measured by the PISA index of economic, social and cultural status), gender, the number of science courses attended and the grade in which the student is enrolled.
Note: Statistically significant values are marked in a darker tone.
Countries and economies are ranked in descending order of the change in science performance associated with a one-unit increase on the index of teacher feedback in science lessons reported by students, after accounting for student characteristics and observed and unobserved school features.
Source: OECD, PISA 2015 Database.

The same analyses were repeated for all science subscales – those based on science competencies, knowledge categories and content areas – with similar results for all of them (the scales are discussed in detail in Annex A). On average across OECD countries, students score either six or seven points lower in science per one-unit increase in the index of TFSL, after accounting for students' characteristics and school effects.[10]

Does teacher feedback really have a negative effect on science performance?

In almost every school system, teacher feedback and science performance are negatively associated. One could be tempted to conclude that teacher feedback hinders student learning. However, given the large body of literature showing the potential benefits of feedback (Hattie and Timperley, 2007) and teachers' predisposition to provide more feedback to struggling students than to outstanding students (Beaman, Wheldall and Kemp, 2006), the negative association between teacher feedback and science performance presumably reflects the greater likelihood that struggling students attract more attention from teachers. Several findings point in this direction, suggesting that teacher feedback may not really have a negative effect on student learning.

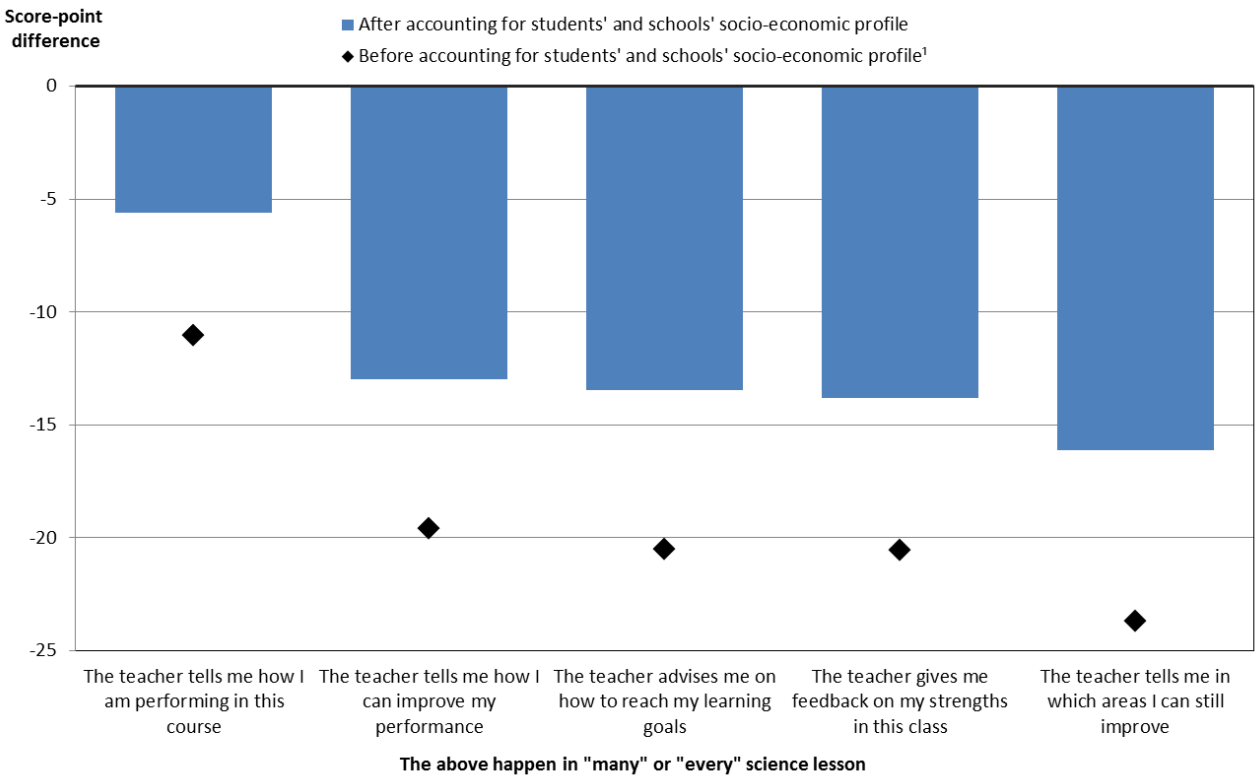
The first of these findings is that, on average across OECD countries and in most school systems, the negative association weakens substantially after accounting for the socio-economic profile of students and schools, which is a strong predictor of student learning

(OECD, 2016b). In the Netherlands, for instance, students score 10 points lower in science per one-unit increase in the index of teacher feedback after accounting for the socio-economic profile of students and schools, compared to 20 points lower before accounting for socio-economic status.

Analysing the relationship with science performance for all the items of the index of teacher feedback, one also observes that the negative association is weakest for the most neutral type of feedback – "the teacher tells me how I am performing in this course" – and strongest for a kind of feedback that is most relevant to underachieving students – "the teacher tell me in which areas I can still improve" (Figure 6.8). Students who reported that their science teacher tells them how they are performing in the course in many lessons or every lesson score only six points lower than students who reported that this never happened, or happened only in some lessons, after accounting for socio-economic status. The same comparison for the response "the teacher tells me how can I improve my performance" yields a difference of 13 score points, while the comparison for the response "the teacher tells me in which areas I can still improve" results in a difference of 16 score points between students who reported that they receive this type of feedback in many or every lesson compared with students who reported that this feedback is never, or only sometimes, given.

Figure 6.8. Teacher feedback practices and science performance

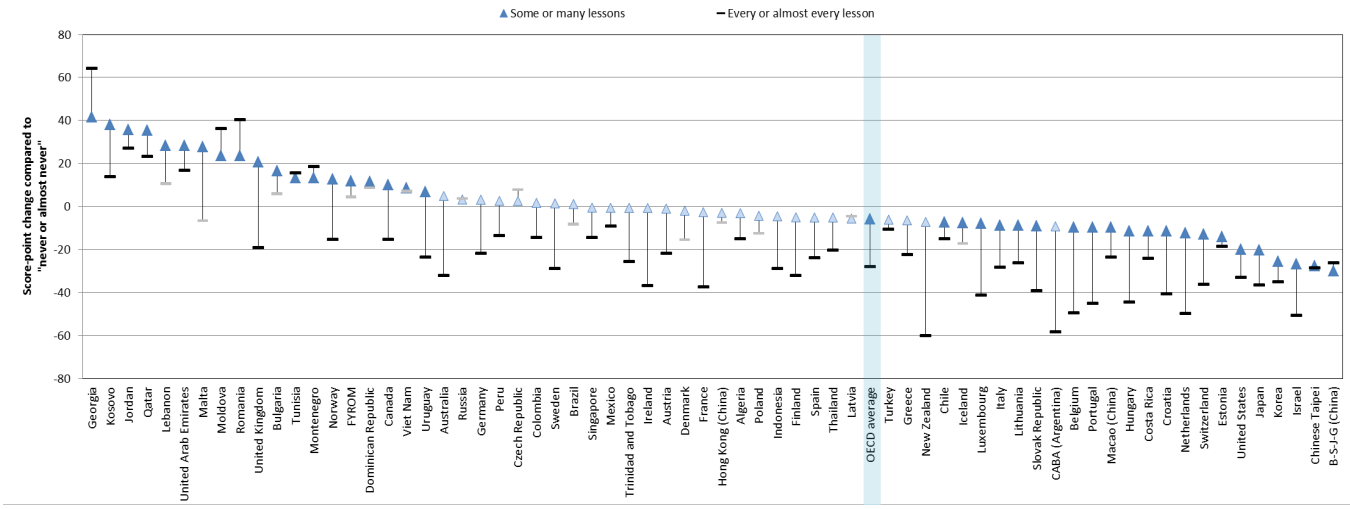
Results based on students' self-reports (OECD average)



1. The socio-economic profile is measured by the PISA index of economic, social and cultural status.
Note: All differences are statistically significant.
Source: OECD, PISA 2015 Database.

Another interesting finding is that only when students receive constant feedback – that is, the type of feedback typically received by struggling and disruptive students (Beaman, Wheldall and Kemp, 2006) – do they score considerably lower in science. For instance, on average across OECD countries, students who reported that their science teacher never tells them how they are performing in the course score 28 points higher than students who reported that this happens in every, or almost every, science lesson, but only 6 points higher than students who reported that this occurs in some or many science lessons (Figure 6.9). In fact, in a large number of school systems, students receiving this type of feedback in some or many science lessons did not score lower than students who reported that they never receive this type of feedback.

Figure 6.9. Telling students how they are performing in the science course and science performance

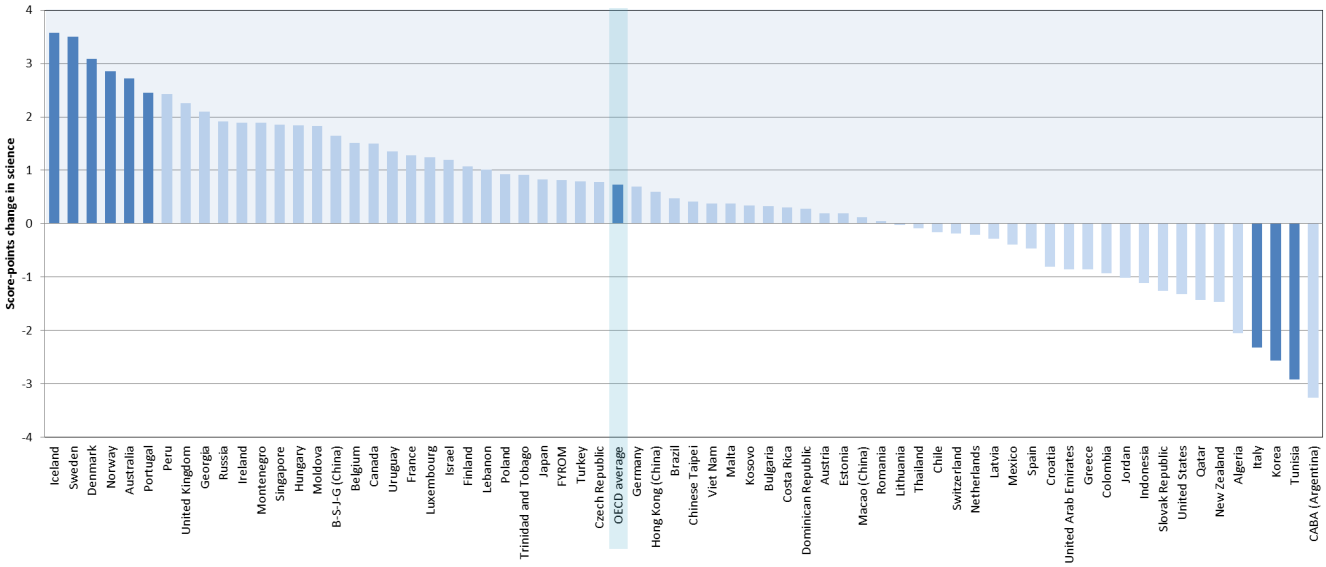


Note: Statistically significant changes are shown in a darker tone.
Countries and economies are ranked in descending order of the score-point difference between students who reported that their science teacher tells them how they are performing in the course in "some of many lessons" and those who reported that their teacher "never or almost never" does.
Source: OECD, PISA 2015 Database.

The last of the findings that mediates the observed negative association between teacher feedback and science performance is that, once performance in reading and mathematics is accounted for, teacher feedback and science performance are positively associated, even if moderately so, on average across OECD countries (Figure 6.10). In other words, when students with similar general academic performance are compared – in this way adjusting for all the student characteristics that explain their overall academic performance, such as general intelligence, academic engagement, self-efficacy and test-taking effort – those who reported receiving more frequent teacher feedback in science lessons score somewhat higher in science than those who reported receiving less frequent feedback.

Figure 6.10. Index of teacher feedback and adjusted science performance

Regression coefficients after accounting for reading and mathematics scores



Note: Statistically significant changes are shown in a darker tone.
Countries and economies are ranked in descending order of the score-point difference between students who reported that their science teacher tells them how they are performing in the course in "some or many lessons" and those who reported that their teacher "never or almost never" does.
Source: OECD, PISA 2015 Database.

Box 6.2. Do smaller classes allow for certain teaching practices more than larger classes?

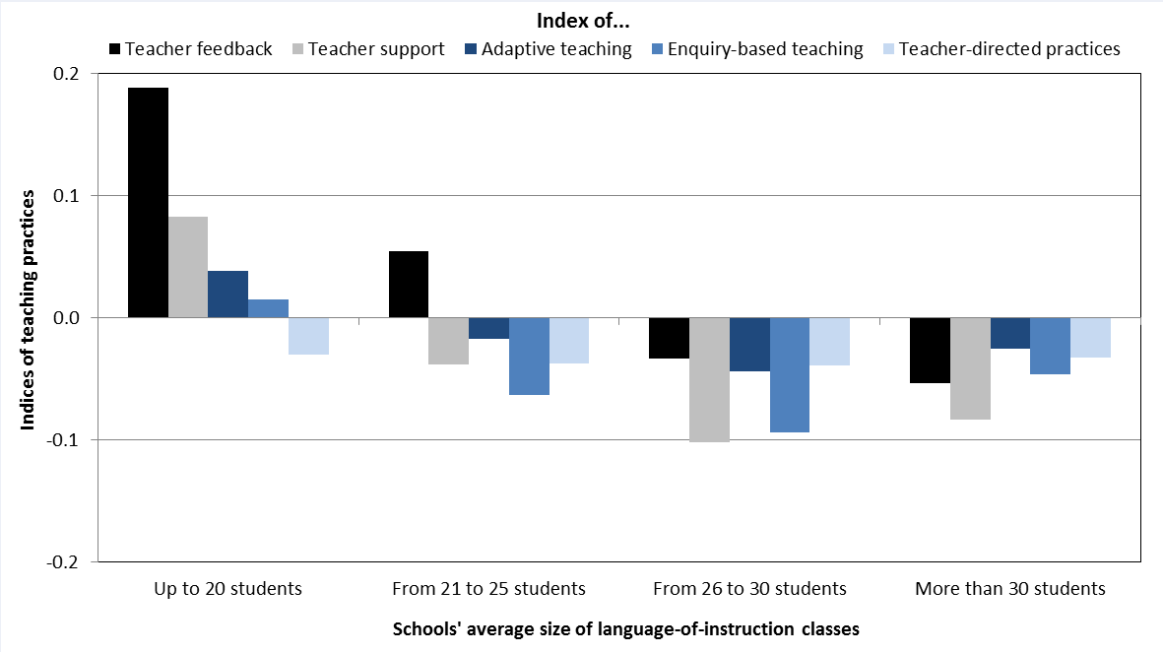
Teachers often complain that certain practices can only be effectively implemented in small classes. Supporting this claim, some research documents how classroom activities and assessments are more varied, active and creative, and instruction more personalised in small classes (Bourke, 1986; Glass and Smith, 1979; Harfitt, 2012), while traditional lectures happen more frequently in large classes (Bourke, 1986). However, other studies show that many teachers do not necessarily allocate instruction time differently (Rice, 1999) or use different practices (Blatchford and Mortimore, 1994) when they are transferred from a large to a small class, particularly if the move is not accompanied by changes in other aspects of teaching, such as the curriculum, assessment practices or teachers' professional development (Evertson and Randolph, 1989; Stecher et al., 2001; Zahorik et al., 2003).

Does the frequency of teaching practices vary across schools with different class sizes? To answer this question, PISA asked school principals about the average size of the language-of-instruction classes in their schools,[11] and asked students about the frequency with which certain teaching practices were used in their (science) classes. Both pieces of data were then combined for the OECD countries with enough variation in class sizes.

According to students, all but teacher-directed practices are more frequently used in schools with small classes ("up to 20 students") than they are in schools with large classes ("more than 30 students") (Figure 6.11). The largest difference is observed for teacher feedback, and to a lesser extent for teacher support, whose frequency decreases steadily with the number of students in language-of-instruction classes. This should come as no surprise given that providing feedback and supporting students require spending more time working with individual students. Adaptive and enquiry-based teaching practices are also more frequently used in smaller classes ("up to 20 students") than in larger classes; but, once the average size of language-of-instruction classes exceeds 20 students, the frequency remains relatively constant.

Figure 6.11. Teaching practices by class size

OECD average

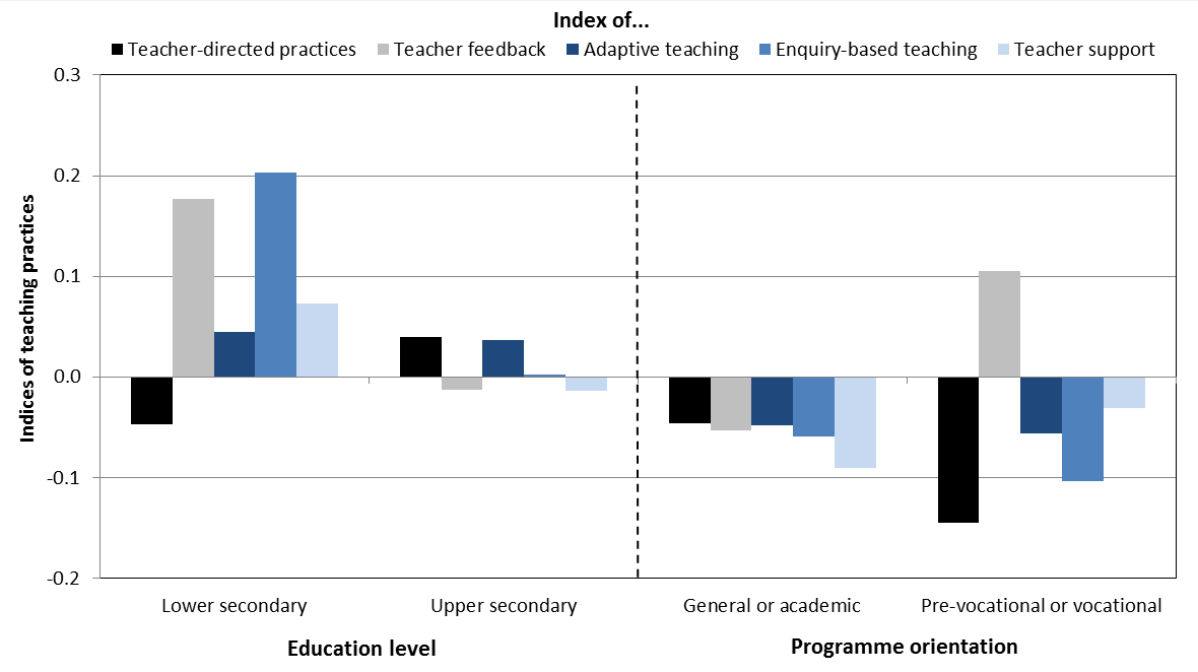


Note: Results are based on the 21 OECD countries with at least 30 students and five schools in each category.
Source: OECD, PISA 2015 Database.

These results may be partly explained by the different teaching approaches used in lower and upper secondary classes. In lower secondary classes, which are typically smaller than upper secondary classes (OECD, 2016b), teacher feedback, enquiry-based teaching and teacher support occur more frequently than in upper secondary classes, according to students' reports (Figure 6.12). Teaching approaches change considerably less between general and vocational programmes; teacher-directed practices are somewhat more frequently used in general programmes, whereas teacher feedback occurs more frequently in pre-vocational or vocational programmes.

Figure 6.12. Teaching practices by educational level and programme orientation

OECD average



Note: The results are based on 27 OECD countries for education level and 20 OECD countries from programme orientation.
Source: OECD, PISA 2015 Database.

7. What the results imply for policy and practice

Much ink has been spilled debating the merits of different teaching practices. The debate, anchored in the philosophical discussion about the nature of learning, has often sparked emotional reactions on the different sides of the argument. Proponents of enquiry-based science teaching argue that this approach allows students to learn through exposure to the procedures of professional scientists, while the proponents of traditional methods emphasise the role of teachers in transmitting knowledge and in guiding enquiry. These two approaches are compounded by rising demands for adaptation of the content of lessons in order to meet the needs of diverse student populations. As such, the reality of teaching in the 21st century is complex as teachers have to rely on different approaches while adapting to rapidly changing educational environments.

In this paper, PISA 2015 data were used to investigate the association between different teaching practices and student cognitive and non-cognitive outcomes. The focus is on enquiry-based science teaching (EBST), teacher-directed science instruction (TDSI), adaptive teaching in science lessons (ATSL), teacher support in science lessons (TSSL), and teacher feedback in science lessons (TFSL). The outcomes of interest are overall science performance and performance in science sub-domains in addition to science-related dispositions, such as interest in and enjoyment of science, science epistemic beliefs, science self-efficacy, and expectations of working in a science career at the age of 30. The following summarises the implications of the results.

Choose teaching strategies suitable for the school context.

The findings show that the frequency of students' exposure to the various teaching strategies varies with the context of the schools and their students. For instance, students report frequent exposure to enquiry-based teaching in disadvantaged and rural schools while they report frequent exposure to teacher-directed instruction in advantaged schools. By contrast, adaptive teaching is more frequently encountered in private schools and teacher feedback in disadvantaged schools. Moreover, all teaching strategies are more frequently used in smaller classes, those with fewer than 20 students. The only exception is teacher-directed instruction, where the frequency of use, as reported by the students, is not related to class size.

These cross-tabulations show that certain practices might be better adapted to particular contexts. For instance, students in disadvantaged schools are likely to be low performers and require more teacher support and feedback. Teaching practices that require intensive interactions between students and teachers are more likely to take place in smaller classes where these discussions can be easily managed. In contrast, teacher-directed instruction is not sensitive to class size and can be used in various contexts.

Foster discipline to ensure the success of enquiry-based science teaching.

The findings of this paper emphasise the role of a positive school climate in the success of enquiry-based science teaching. The negative association between EBST and science achievements, exacerbated in science classes lacking discipline, is attenuated when students attend disciplined science lessons. This finding highlights the sensitivity of this practice to the surrounding environment and explains, to a certain extent, the negative

correlation between EBST and science performance. The lack of discipline in a classroom challenges teachers' ability to organise EBST activities and, in turn, reduces the effectiveness of this practice.

The negative association between EBST and science performance is not found to vary by student performance or by some student and school characteristics, such as interest in and enjoyment of science, and school resources. However, the negative association between EBST and success in answering science items correctly is found to be weaker for difficult items than for easy items.

In spite of the negative association between EBST and science performance, exposure to science-enquiry activities is positively associated with various non-cognitive science outcomes and attitudes, such as enjoyment of and interest in science, epistemic beliefs, science self-efficacy, and career expectations. This positive association holds true even after accounting for science performance, and student and school characteristics. More interestingly, this study implies that exposing all students, particularly girls, to enquiry-based activities, like designing experiments, may improve their attitudes towards science and raise their expectations of pursuing a career in this field. In fact, exposing girls to these teaching approaches may help close the gender gap in traditionally male-dominated fields, such as engineering and science-related technical professions. On average across OECD countries, the probability of expecting a career in engineering and other male-dominated occupations increases more among girls than boys when students are exposed to enquiry-based teaching practices.

Teacher-directed science instruction is always a reliable strategy regardless of the school context.

Teacher-directed science instruction is positively associated with science performance in almost all countries, even after accounting for student characteristics, and observed and unobserved school features. This positive association is not found to vary with performance in science sub-domains, the level of performance, the difficulty of science test questions, or student and school characteristics (e.g. disciplinary climate, student composition, resources, etc.). As such, TDSI is a robust teaching practice that is likely to deliver on its promise regardless of the surrounding environment.

However, the findings show no evidence that the joint use of EBST and TDSI is positively associated with science performance. In other words, combining EBST activities with teacher-directed instruction (e.g. teacher tutoring, scaffolded discovery) is not necessarily beneficial.

Use adaptive teaching and teacher support to enhance the learning experience of diverse student populations.

Findings show that adaptive teaching in science lessons is positively correlated with science performance in the majority of countries, even after accounting for student and school characteristics. This relationship is particularly strong in the Nordic countries, which are known for their comprehensive education systems and for their reliance on personalised learning approaches. This positive association does not vary by performance in science sub-domains or by students' level of performance.

Teacher support in science lessons was found to be positively and significantly correlated with science performance in more than 30 of the participating countries and economies.

However, the association is weak. As such, teacher support could be seen as a complementary strategy that could reinforce other practices by remediating their shortcomings in dealing with the needs of a diverse student population.

Teacher feedback complements other strategies and could help the low performers.

The negative association between teacher feedback and students' science performance can be attributed to the fact that students who require the most feedback tend to be low performers. This negative association holds even after accounting for student characteristics and school features. However, when the frequency of feedback is taken into account, the findings show that it is the most frequent feedback that is negatively associated with performance. Moreover, after adjusting for performance in other subjects (reading and mathematics), the association between feedback and science performance becomes positive, albeit weak, across OECD countries. This indicates that feedback could be beneficial once students' underlying abilities have been accounted for.

All teaching practices are effective in improving students' attitudes and dispositions towards science.

In addition to science performance, the analyses focus on students' non-cognitive outcomes, such as interest in and enjoyment of science, epistemic beliefs, science self-efficacy and career expectations. The findings show a clear, positive association between these outcomes and teaching strategies, even after adjusting for students' science performance, and for student and school characteristics. The associations are significant in most countries, especially for interest in and enjoyment of science, epistemic beliefs, and science self-efficacy, though less so for career expectations.

The results show that different teaching practices used in different contexts could foster positive attitudes and dispositions towards science. Some of these practices could be combined to enhance student learning.

References

Aaronson, D., L. Barrow, and W. Sander (2007), "Teachers and student achievement in the Chicago public high schools", *Journal of Labor Economics*, Vol. 25/1, pp. 95–135, http://dx.doi.org/10.1086/508733.

Bandura, A. (1997), *Self-Efficacy: The Exercise of Control*, W.H. Freeman, New York.

Beaman, R., K. Wheldall and C. Kemp (2006), "Differential teacher attention to boys and girls in the classroom", *Educational Review*, Vol. 58/3, pp. 339–366, http://dx.doi.org/10.1080/00131910600748406.

Black, P. and D. Wiliam (1998), "Assessment and classroom learning", *Assessment in Education: Principles, Policy & Practice*, Vol. 5/1, pp. 7–74, http://dx.doi.org/10.1080/0969595980050102.

Blatchford, P. and P. Mortimore (1994), "The issue of class size for young children in schools: What can we learn from research?", *Oxford Review of Education*, Vol. 20/4, pp. 411–428, http://dx.doi.org/10.1080/0305498940200402.

Borich, G. (2011), *Effective teaching methods : research-based practice*, Pearson/Allyn & Bacon.

Bourke, S. (1986), "How smaller is better: Some relationships between class size, teaching practices, and student achievement", *American Educational Research Journal*, Vol. 23/4, pp. 558–571, http://dx.doi.org/10.3102/00028312023004558.

Bulman, G. and R. Fairlie (2016), "Technology and education: Computers, software, and the Internet." In E. A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of the Economics of Education*, vol. 5, pp. 239–280, https://doi.org/10.1016/B978-0-444-63459-7.00005-1.

Butler, A., J. Karpicke and H. Roediger (2007), "The effect of type and timing of feedback on learning from multiple-choice tests.", *Journal of Experimental Psychology: Applied*, Vol. 13/4, pp. 273–281, http://dx.doi.org/10.1037/1076-898X.13.4.273.

Cakir, M. (2008), "Constructivist approaches to learning in science and their implications for science pedagogy: A literature review", *International Journal of Environmental & Science Education*, Vol. 3/4, pp. 193–206, https://files.eric.ed.gov/fulltext/EJ894860.pdf.

Chen, L. et al. (2017), *Drivers of Student Performance: Insights from Asia*, McKinsey and company.

Cook, S. D. N. and J. S. Brown (1999), "Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing", *Organization Science*, Vol. 10/4, pp. 381–400, https://doi.org/10.1287/orsc.10.4.381.

Crawford, B. A. (2007), "Learning to teach science as inquiry in the rough and tumble of practice", *Journal of Research in Science Teaching*, Vol. 44/4, pp. 613–642, https://doi.org/10.1002/tea.20157.

Cunningham, D. J. (1991), "Assessing constructions and constructing assessments: A dialogue", *Educational Technology*, Vol. 31/5, pp. 13–17, https://www.jstor.org/stable/44427515.

de Groot, A. D. (1969), *Methodology: Foundations of Inference and Research in the Behavioral Sciences*, Mouton, The Hague.

Dee, T. S. (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.

Dee, T. S. (2007), "Teachers and the gender gaps in student achievement", *Journal of Human Resources*, Vol. XLII/3, pp. 528–554, http://dx.doi.org/10.3368/jhr.XLII.3.528.

Dochy, F. et al. (2003), "Effects of problem-based learning: a meta-analysis", *Learning and Instruction*, Vol. 13/5, pp. 533–568, https://doi.org/10.1016/S0959-4752(02)00025-7.

Dunbar, R. and M. O'Sullivan (1986), "Effects of intervention on differential treatment of boys and girls in elementary physical education lessons", *Journal of Teaching in Physical Education*, Vol. 5/3, pp. 166–175, http://dx.doi.org/10.1123/jtpe.5.3.166.

Duncombe, W. and J. Yinger (2007), "Does school district consolidation cut costs?", *Education Finance and Policy*, Vol. 2/4, pp. 341–375, http://dx.doi.org/10.1162/edfp.2007.2.4.341.

Dupriez, V., X. Dumay, and A. Vause (2008), "How do school systems manage pupils' heterogeneity?", *Comparative Education Review*, Vol. 52/2, pp. 245–273, https://doi.org/10.1086/528764.

Duschl, R. A. et al. (eds.) (2007), *Taking Science to School: Learning and Teaching Science in Grades K-8*, National Academies Press, Washington, DC.

Dweck, C. et al. (1978), "Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis.", *Developmental Psychology*, Vol. 14/3, pp. 268–276, http://dx.doi.org/10.1037/0012-1649.14.3.268.

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

Evertson, C. and C. Randolph (1989), "Teaching practices and class size: A new look at an old issue", *Peabody Journal of Education*, Vol. 67/1, pp. 85–105, https://doi.org/10.1080/01619569209538671.

Eylon, B. and M. Linn (1988), "Learning and instruction: An examination of four research perspectives in science education", *Review of Educational Research*, Vol. 58/3, pp. 251–301, http://dx.doi.org/10.2307/1170256.

Falck, O., C. Mang, and L. Woessman (2018), "Virtually no effect? Different uses of classroom computers and their effects on student achievement." *Oxford Bulletin of Economics and Statistics*, Vol. 80/1, pp. 1–38. https://doi.org/10.1111/obes.12192.

Fayer, S., A. Lacey, and A. Watson (2017), "STEM Occupations: Past, Present, And Future", *Spotlight on Statistics*, U.S. Department of Labor, Bureau of Labor Statistics, Washington, DC.

Federici, R. A. and E. M. Skaalvik (2014), "Students' perceptions of emotional and instrumental teacher support: relations with motivational and emotional responses", *International Education Studies*, Vol. 7/1, pp. 21–36, http://dx.doi.org/10.5539/ies.v7n1p21.

Feldman, M. S. and B. T. Pentland (2003), "Reconceptualizing organizational routines as a source of flexibility and change", *Administrative Science Quarterly*, Vol. 48/1, pp. 94–118, https://doi.org/10.2307%2F3556620.

Ferguson-Hessler, M. G. M. and T. De Jong (1990), "Studying physics texts: differences in study processes between good and poor performers", *Cognition and Instruction*, Vol. 7/1, pp. 41–54, https://doi.org/10.1207/s1532690xci0701_2.

French, J. and P. French (1984), "Gender imbalances in the primary classroom: an interactional account", *Educational Research*, Vol. 26/2, pp. 127–136, http://dx.doi.org/10.1080/0013188840260209.

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

Furtak, E. M. et al. (2012), "Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis", *Review of Educational Research*, Vol. 82/3, pp. 300–329, https://doi.org/10.3102%2F0034654312457206.

Gago, J. et al. (2004), *Increasing human resources for science and technology in Europe*, European Commission, Brussels.

Gallimore, R. et al. (2009), "Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams", *The Elementary School Journal*, Vol. 109/5, pp. 537–553, https://doi.org/10.1086/597001.

Galton, M. (2009), "Moving to secondary school: Initial encounters and their effects", *Perspectives on education*, Vol. 2, pp. 5-21, https://www.researchgate.net/publication/258984426 (accessed on 06 June 2018).

Gardner, H. (2011), *Frames of Mind: The Theory of Multiple Intelligences*, Basic Books, New York.

Geier, R. et al. (2008), "Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform", *Journal of Research in Science Teaching*, Vol. 45/8, pp. 922–939, https://doi.org/10.1002/tea.20248.

Glass, G. and M. Smith (1979), "Meta-analysis of research on class size and achievement", *Educational Evaluation and Policy Analysis*, Vol. 1/1, pp.2–16, http://dx.doi.org/10.2307/1164099.

Guess, P. E. and S. J. McCane-Bowling (2016), "Teacher support and life satisfaction: An investigation with urban, middle school students", *Education and Urban Society*, Vol. 48/1, pp. 30–47, https://doi.org/10.1177%2F0013124513514604.

Handelsman, J. et al. (2004), "EDUCATION: Scientific Teaching", *Science*, Vol. 304/5670, pp. 521–522, http://dx.doi.org/10.1126/science.1096022.

Harfitt, G. (2012), "An examination of teachers' perceptions and practice when teaching large and reduced-size classes: Do teachers really teach them in the same way?", *Teaching and Teacher Education*, Vol. 28/1, pp. 132–140, http://dx.doi.org/10.1016/J.TATE.2011.09.001.

Hattie, J. (2009), *Visible learning : a synthesis of over 800 meta-analyses relating to achievement*, Routledge, New York and London.

Hattie, J., M. Gan and C. Brooks (2017), "Instruction based on feedback", in Mayer, R. and P. Alexander (eds.), *Handbook of Research on Learning and Instruction*, Routledge, New York and London.

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.

Hickey, D. T., E. W. Wolfe, and A. C. H. Kindfield (2000), "Assessing learning in a technology-supported genetics environment: Evidential and systemic validity issues", *Educational Assessment*, Vol. 6/3, pp. 155–196, https://doi.org/10.1207/S15326977EA0603_1.

Hofstein, A. and V. Lunetta (2004), "The laboratory in science education: Foundations for the twenty-first century", *Science Education*, Vol. 88/1, pp. 28–54, http://dx.doi.org/10.1002/sce.10106.

Hmelo-Silver, C. E., R. G. Duncan, and C. A. Chinn (2007), "Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006)", *Educational Psychologist*, Vol. 42/2, pp. 99–107, https://doi.org/10.1080/00461520701263368.

Hofer, B. K. and P. R. Pintrich (1997), "The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning", *Review of Educational Research*, Vol. 67/1, pp. 88–140, https://doi.org/10.3102%2F00346543067001088.

Hutchinson, J. (2012), "Career-related learning and science education: The changing landscape", *School Science Review*, No. 346, pp. 91–98, http://hdl.handle.net/10545/243597.

Hyland, F. (2003), "Focusing on form: student engagement with teacher feedback", *System*, Vol. 31/2, pp. 217–230, http://dx.doi.org/10.1016/S0346-251X(03)00021-6.

Ikwumelu, S., O. Oyibe and E. Oketa (2015), "Adaptive teaching: An invaluable pedagogic practice in social studies education", *Journal of Education and Practice*, Vol. 6/33, pp.140–144, https://files.eric.ed.gov/fulltext/EJ1083517.pdf (accessed on 26 April 2018).

Jonassen, D. H. (1991), "Objectivism versus constructivism: Do we need a new philosophical paradigm?", *Educational Technology Research and Development*, Vol. 39/3, pp. 5–14, https://www.jstor.org/stable/30219973.

Jones, S. M., and Dindia, K. (2004), "A meta-analytic perspective on sex equity in the classroom" *Review of educational research*, 74/4, pp. 443-471, http://journals.sagepub.com/doi/pdf/10.3102/00346543074004443.

Kane, T. J., J. E. Rockoff, and D. O. 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, https://doi.org/10.1016/j.econedurev.2007.05.005.

Kelly, A. (1988), “Gender differences in teacher–pupil interactions: a meta-analytic review”, *Research in Education*, Vol. 39/1, pp. 1–23, http://dx.doi.org/10.1177/003452378803900101.

Kirschner, P. A. (1992), "Epistemology, practical work and academic skills in science education", *Science and Education*, Vol. 1/3, pp. 273–299, https://doi.org/10.1007/BF00430277.

Kirschner, P. A., J. Sweller, and R. E. Clark (2006), "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching", *Educational Psychologist*, Vol. 41/2, pp. 75–86, https://doi.org/10.1207/s15326985ep4102_1.

Klem, A. M. and J. P. Connell (2004), "Relationships matter: Linking teacher support to student engagement and achievement", *Journal of School Health*, Vol. 74/7, pp. 262–273, https://doi.org/10.1111/j.1746-1561.2004.tb08283.x.

Kloser, M. (2014), "Identifying a core set of science teaching practices: A delphi expert panel approach: CORE SCIENCE TEACHING PRACTICES", *Journal of Research in Science Teaching*, Vol. 51(9), pp. 1185–1217, https://doi.org/10.1002/tea.21171.

Kluger, A. and A. DeNisi (1996), “The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory.”, *Psychological Bulletin*, Vol. 119/2, pp. 254–284, http://dx.doi.org/10.1037/0033-2909.119.2.254.

Konstantopoulos, S. and V. Chung (2009), “What are the long-term effects of small classes on the achievement gap? Evidence from the lasting benefits study”, *American Journal of Education*, Vol. 116/1, pp. 125–154, http://dx.doi.org/10.1086/605103.

Krajcik, J. S., C. Czerniak, and C. Berger (1999), *Teaching Children Science: A Project-Based Approach*, McGraw-Hill, Boston.

Kulhavy, R. (1977), “Feedback in Written Instruction”, *Review of Educational Research*, Vol. 47/2, pp. 211–232, http://dx.doi.org/10.3102/00346543047002211.

Lee, I. (2008), “Student reactions to teacher feedback in two Hong Kong secondary classrooms”, *Journal of Second Language Writing*, Vol. 17/3, pp. 144–164, http://dx.doi.org/10.1016/J.JSLW.2007.12.001.

Loewenberg Ball, D. and F. M. Forzani (2009), "The work of teaching and the challenge for teacher education", *Journal of Teacher Education*, Vol. 60/5, pp. 497–511, https://doi.org/10.1177%2F0022487109348479.

Lynch, S. et al. (2005), "Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant", *Journal of Research in Science Teaching*, Vol. 42/8, pp. 912–946, https://doi.org/10.1002/tea.20080.

Mason, L. et al. (2013), "Besides knowledge: a cross-sectional study on the relations between epistemic beliefs, achievement goals, self-beliefs, and achievement in science", *Instructional Science*, Vol. 41/1, pp. 49–79, https://doi.org/10.1007/s11251-012-9210-0.

Mayer, R. E. (2004), "Should there be a three-strikes rule against pure discovery learning?", *American Psychologist*, Vol. 59/1, pp. 14–19, https://doi.org/10.1037/0003-066X.59.1.14.

McGinnis, J. R., C. Parker and A. O. Graeber (2004), "A cultural perspective of the induction of five reform-minded beginning mathematics and science teachers", *Journal of Research in Science Teaching*, Vol. 41/7, pp. 720–747, https://doi.org/10.1002/tea.20022.

Minner, D. D., A. J. Levy and J. Century (2010), "Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002", *Journal of Research in Science Teaching*, Vol. 47/4, pp. 474–496, https://doi.org/10.1002/tea.20347.

Mostafa, T. and J. Pál (2018), "Science teachers' satisfaction", *OECD Education Working Papers* No. 168, OECD Publishing, Paris, https://doi.org/10.1787/1ecdb4e3-en.

National Research Council (U.S.) (Ed.), (1996) *National Science Education Standards: Observe, Interact, Change, Learn*, National Academy Press, Washington, DC.

Newman, W. J. et al. (2004), "Dilemmas of teaching inquiry in elementary science methods", *Journal of Science Teacher Education*, Vol. 15/4, pp. 257–279, https://doi.org/10.1023/B:JSTE.0000048330.07586.d6.

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

OECD (2017a), *PISA 2015 Technical Report*, OECD Publishing, Paris, http://www.oecd.org/pisa/data/2015-technical-report/ (accessed on 08 November 2018)

OECD (2017b), *PISA 2015 Results (Volume V): Collaborative Problem Solving*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264285521-en.

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

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

OECD (2008), *Encouraging Student Interest in Science and Technology Studies*, OECD Publishing, Paris, https://doi.org/10.1787/9789264040892-en.

Olson, S. and D. G. Riordan (2012), *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President*, Executive Office of the President.

Opitz, B., N. Ferdinand and A. Mecklinger (2011), "Timing matters: the impact of immediate and delayed feedback on artificial language learning.", *Frontiers in human neuroscience*, Vol. 5, p. 8, http://dx.doi.org/10.3389/fnhum.2011.00008.

Ormrod, J. E. (2012), *Essentials of Educational Psychology: Big Ideas to Guide Effective Teaching*, 3rd ed., Pearson, Boston.

Osborne, J., S. Simon and S. Collins (2003), "Attitudes towards science: A review of the literature and its implications", *International Journal of Science Education*, Vol. 25/9, pp. 1049–1079, https://doi.org/10.1080/0950069032000032199.

Pitzer, J. and E. Skinner (2017), "Predictors of changes in students' motivational resilience over the school year: The roles of teacher support, self-appraisals, and emotional reactivity", *International Journal of Behavioral Development*, Vol. 41/1, pp. 15–29, https://doi.org/10.1177%2F0165025416642051.

Reckwitz, A. (2002), "Toward a theory of social practices: A development in culturalist theorizing", *European Journal of Social Theory*, Vol. 5/2, pp. 243–263, https://doi.org/10.1177%2F13684310222225432.

Reinke, W., K. Herman and L. Newcomer (2016), "The brief student–teacher classroom interaction observation", *Assessment for Effective Intervention*, Vol. 42/1, pp. 32–42, http://dx.doi.org/10.1177/1534508416641605.

Ricard, N. C. and L. G. Pelletier (2016), "Dropping out of high school: The role of parent and teacher self-determination support, reciprocal friendships and academic motivation", *Contemporary Educational Psychology*, Vol. 44–45, pp. 32–40, https://doi.org/10.1016/j.cedpsych.2015.12.003.

Rice, J. (1999), "The impact of class size on instructional strategies and the use of time in high school mathematics and science courses", *Educational Evaluation and Policy Analysis*, Vol. 21/2, pp. 215–229, http://dx.doi.org/10.3102/01623737021002215.

Rieber, L. P. and M. W. Parmley (1995), "To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science", *Journal of Educational Computing Research*, Vol. 13/4, pp. 359–374, https://doi.org/10.2190%2FM8VX-68BC-1TU2-B6DV.

Rivkin, S. G., E. A. Hanushek and J. F. Kain (2005), "Teachers, schools, and academic achievement", *Econometrica*, Vol. 73/2, pp. 417–458, https://doi.org/10.1111/j.1468-0262.2005.00584.x.

Rockoff, J. E. (2004), "The impact of individual teachers on student achievement: Evidence from panel data", *American Economic Review*, Vol. 94/2, pp. 247–252, http://www.jstor.org/stable/3592891.

Rodrigues, M., and F. Biagi (2017), "Digital technologies and learning outcomes of students from low socio-economic background: An Analysis of PISA 2015", Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2760/415251.

Sadler, D. (1989), "Formative assessment and the design of instructional systems", *Instructional Science*, Vol. 18/2, pp. 119–144, http://dx.doi.org/10.1007/BF00117714.

Schafft, K. (2016), "Rural education as rural development: Understanding the rural school–community well-being linkage in a 21st-century policy context", *Peabody Journal of Education*, Vol. 91/2, pp. 137–154, http://dx.doi.org/10.1080/0161956X.2016.1151734.

Schwartz, A. (2006), "Contextualized chemistry education: The American experience", *International Journal of Science Education*, Vol. 28/9, pp. 977–998, http://dx.doi.org/10.1080/09500690600702488.

Schwartz, D. L. and J. D. Bransford (1998), "A time for telling", *Cognition and Instruction*, Vol. 16/4, pp. 475–522, http://psycnet.apa.org/doi/10.1207/s1532690xci1604_4.

Schwerdt, G. and A. C. Wuppermann (2011), "Is traditional teaching really all that bad? A within-student between-subject approach", *Economics of Education Review*, Vol. 30/2, pp. 365–379, https://doi.org/10.1016/j.econedurev.2010.11.005.

Seidel, T. and R. J. 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, https://doi.org/10.3102%2F0034654307310317.

Shute, V. J. (1990), "A comparison of inductive and deductive learning environments: which is better for whom and why?", Paper presented at the American Educational Research Association (AERA) Annual Meeting, Boston, USA.

Shute, V. J., R. Glaser and K. Raghavan (1989), "Inference and discovery in an exploratory laboratory", in P. L. Ackerman, R. J. Sternberg and R. Glaser (eds.), *Learning and Individual Differences: Advances in Theory and Research*, pp. 279–326, W.H. Freeman & Company, New York.

Spillane, J. P. (2012), "Data in practice: Conceptualizing the data-based decision-making phenomena", *American Journal of Education*, Vol. 118/2, pp. 113–141, https://doi.org/10.1086/663283.

Stecher, B. et al. (2001), "Class-size reduction in California: A story of hope, promise, and unintended consquences", *PHI DELTA KAPPAN*, Vol. 82/9, pp.670-674, http://journals.sagepub.com/doi/pdf/10.1177/003172170108200907 (accessed on 04 June 2018).

Strijbos, J.-W., P. A. Kirschner and R. L. Martens (eds.) (2004), *What We Know About CSCL*, Springer, Dordrecht, Netherlands.

Sully De Luque, M. and S. Sommer (2000), "The impact of culture on feedback-seeking behavior: An integrated model and propositions", *Academy of Management Review*, Vol. 25/4, pp. 829–849, http://dx.doi.org/10.5465/amr.2000.3707736.

Swaak, J., T. De Jong and W. R. Van Joolingen (2004), "The effects of discovery learning and expository instruction on the acquisition of definitional and intuitive knowledge: Discovery learning vs. expository instruction", *Journal of Computer Assisted Learning*, Vol. 20/4, pp. 225–234, https://doi.org/10.1111/j.1365-2729.2004.00092.x.

Sweller, J. (2003), "Evolution of human cognitive architecture", in B. H. Ross (ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 43, pp. 215–266, Elsevier Science, New York, http://dx.doi.org/10.1016/S0079-7421(03)01015-6.

Sweller, J. (2004), "Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture", *Instructional Science*, Vol. 32/1-2, pp. 9–31, https://doi.org/10.1023/B:TRUC.0000021808.72598.4d.

Sweller, J., J. J. G. van Merrienboer and F. G. W. C. Paas (1998), "Cognitive architecture and instructional design", *Educational Psychology Review*, Vol. 10/3, pp. 251–296, https://doi.org/10.1023/A:1022193728205.

Sweller, J. (1999), *Instructional Design*, ACER Press, Camberwell.

Tai, R. H., C. Q. Liu, A. V. Maltese and X. Fan (2006), "Planning early for careers in science", *Life Sci*, Vol. 1, pp. 0–2.

Tunstall, P. and C. Gsipps (1996), "Teacher feedback to young children in formative assessment: A typology", *British Educational Research Journal*, Vol. 22/4, pp. 389–404, http://dx.doi.org/10.1080/0141192960220402.

UNESCO (2016), *Closing the gender gap in STEM: Drawing more girls and women into Science, Technology, Engineering and Mathematics*, http://unesdoc.unesco.org/images/0024/002457/245717e.pdf (accessed on 06 June 2018).

Voerman, L. et al. (2012), "Types and frequencies of feedback interventions in classroom interaction in secondary education", *Teaching and Teacher Education*, Vol. 28/8, pp. 1107–1115, http://dx.doi.org/10.1016/J.TATE.2012.06.006.

Vygotskij, L. S. and M. Cole (1981), *Mind in Society: The Development of Higher Psychological Processes*, Harvard Univ. Press, Cambridge, MA.

Weiss, I. R. (1997), "The status of science and mathematics teaching in the United States: Comparing teacher views and classroom practice to national standards.", *NISE Brief*, Vol. 1/3 .

Willms, D. and X. Ma (2004), "School disciplinary climate: characteristics and effects on eighth grade achievement", *Alberta Journal of Educational Research*, Vol. 50/2, pp. 169–188, https://search.proquest.com/openview/f9e8fb72ff086df70f0be22e06887979/1?pq-origsite=gscholar&cbl=34114 (accessed on 20 April 2018).

Windschitl, M. (2003), "Inquiry projects in science teacher education: What can investigative experiences reveal about teacher thinking and eventual classroom practice?", *Science Education*, Vol. 87/1, pp. 112–143, https://doi.org/10.1002/sce.10044.

Windschitl, M., J. Thompson and M. Braaten (2008), "Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations", *Science Education*, Vol. 92/5, pp. 941–967, https://doi.org/10.1002/sce.20259.

World Bank (2016), *World Development Report 2016: Digital Dividends*, World Bank, Washington, DC.

Zahorik, J. et al. (2003), "Teaching practices for smaller classes", *Educational leadership*, Vol. 61/1, pp.75–77, http://www.asu.edu/educ/epsl/EPRU/documents/EPRU-0309-30-RW.pdf (accessed on 04 June 2018).

Zhang, S. (1995), "Reexamining the affective advantage of peer feedback in the ESL writing class", *Journal of Second Language Writing*, Vol. 4/3, pp. 209–222, http://dx.doi.org/10.1016/1060-3743(95)90010-1.

Notes

[1] Sixteen countries and economies took the paper-based test: Albania, Algeria, Argentina, the Former Yugoslav Republic of Macedonia, Georgia, Indonesia, Jordan, Kazakhstan, Kosovo, Lebanon, Malta, Moldova, Romania, Trinidad and Tobago, Viet Nam and Puerto Rico (US).

[2] Teachers are defined as "those whose primary or major activity in school is student instruction, whether it happens in a classroom, in a small group, on a one-to-one basis, or outside regular classrooms". In order to ensure adequate representation of teachers and to guarantee samples that are sufficiently large, sampling of teachers included teachers who were eligible to teach the modal grade of 15-year-old students – whether they were teaching it currently, had done so before, or will/could do so in the future.

[3] Results are available upon request.

[4] See https://www.unr.edu/education/centers-student-resources/initiatives/girls-math-camp/resources/educators/tips.

[5] The regressions were repeated while including a quadratic term of TDSI in order to investigate the presence of non-linear associations between the index and science performance. The results showed non-linear relationships in some countries (although the nature and magnitude of the relationship varies from one country to another); however the adjusted R-squared did not rise by much.

[6] The analysis combining both teaching practices relies on a regression that takes into account the two indices – EBST and TDSI – and their interaction.

[7] Results are available upon request.

[8] According to (Voerman et al., 2012), progress feedback compares the initial level versus the current performance, whereas discrepancy feedback compares the current level versus the desired level of performance.

[9] Results are available upon request.

[10] Results are available upon request.

[11] Please note that class sizes may be different between language-of-instruction and science classes though the expectation is that both are highly correlated across schools.

Annex A

Science test items

PISA 2015 focused on science as the major domain assessed. It defines science literacy as "the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen". A scientifically literate person is willing to engage in reasoned discourse about science and technology. This requires the competencies to explain phenomena, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

In order to assess science performance, 184 test items were developed and selected for PISA 2015 representing the equivalent of six hours of test questions. Of these items, 85 questions (the equivalent of about three hours) were trend tasks, which were used in previous PISA surveys, and 99 questions (another three hours) were new science tasks.

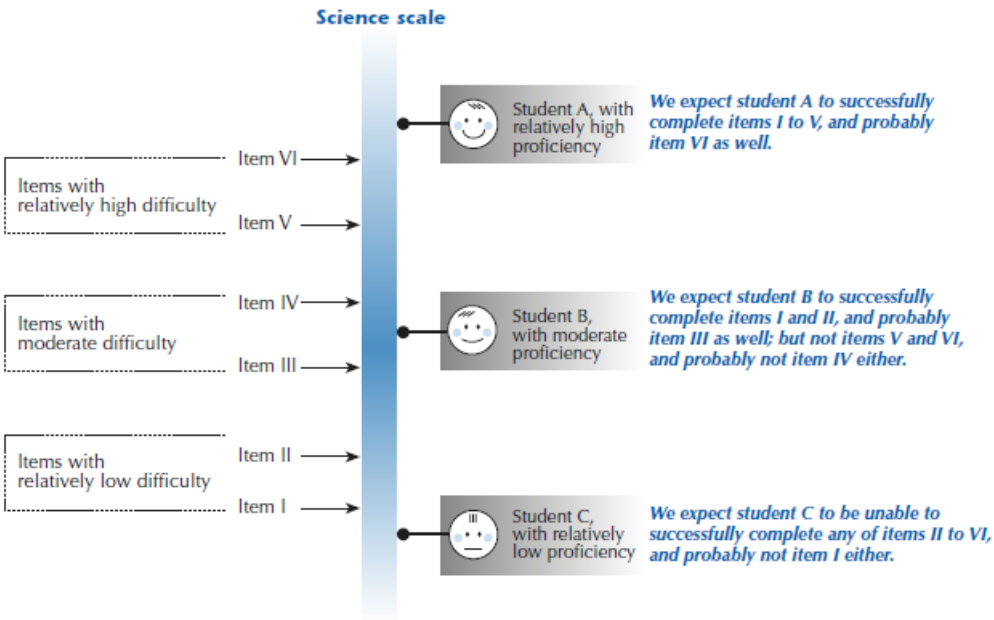
While different students answered different questions, the test design made it possible to construct a continuous scale of proficiency in science, so that each test-taker's performance is associated with a particular point on the scale that indicates his or her estimated science proficiency, and the likelihood that he or she responds correctly to a particular question (higher values on the scale indicate greater proficiency).

The relative difficulty of tasks was estimated by determining the proportion of test-takers who answer each question correctly. Task difficulty is reported on the same scale as student proficiency (higher values correspond, in this case, to more difficult items). In PISA, the difficulty of a task is defined as the point on the scale where there is at least a 62% probability of a correct response by students who score at or above that point. A single continuous scale shows the relationship between the difficulty of questions and the proficiency of test-takers (Figure A.1). By constructing a scale that shows the difficulty of each question, it is possible to locate the level of science literacy that the question demands. By showing the proficiency of each test-taker on the same scale, it is possible to describe each test-taker's level of science literacy. The development of the proficiency measures is described in detail in the PISA 2015 technical report.

Science subscales

Performance in science requires three forms of knowledge: scientific *competencies*, *knowledge* of the standard methodological procedures used in science, and knowledge of science *subject content*. These three domains are interconnected. Explaining scientific and technological phenomena, for instance, demands knowledge of the content of science. Evaluating scientific enquiry and interpreting evidence scientifically also require an understanding of how scientific knowledge is established and the degree of confidence with which it is held.

Figure A.1. Relationship between questions and student performance on a scale



Source: OECD (2016a), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, https://doi.org/10.1787/9789264266490-en.

Scientific competencies subscales

According to the PISA definition, a science-literate person is able and willing to engage in reasoned discourse about science and technology. This requires the competencies to:

- *explain phenomena scientifically*: recognise, offer and evaluate explanations for a range of natural and technological phenomena
- *evaluate and design scientific enquiry*: describe and appraise scientific investigations and propose ways of addressing questions scientifically
- *interpret data and evidence scientifically*: analyse and evaluate data, claims and arguments in a variety of representations and draw appropriate scientific conclusions.

Knowledge categories subscales

Each of the scientific competencies requires some content knowledge (knowledge of theories, explanatory ideas, information and facts), but also an understanding of how such knowledge has been derived (procedural knowledge) and of the nature of that knowledge (epistemic knowledge).

- *Procedural knowledge* refers to knowledge about the concepts and procedures that are essential for scientific enquiry, and that underpin the collection, analysis and interpretation of scientific data. In the quest to explain phenomena in the material world, science proceeds by testing hypotheses through empirical enquiry. Empirical enquiry relies on certain standard procedures to obtain valid and reliable data. Students are expected to know these procedures and related concepts, such as: the notion of dependent and independent variables; the distinction between different types of measurement (qualitative and quantitative, categorical and continuous); ways of assessing and minimising uncertainty (such as repeating measurements); the strategy of controlling variables and its role in

experimental design; and common ways of presenting data. It is expected, for instance, that students will know that scientific knowledge is associated with differing degrees of certainty, depending on the nature and quantity of empirical evidence that has accumulated over time.

- *Epistemic knowledge* refers to an understanding of the nature and origin of knowledge in science, and reflects students' capacity to think and engage in reasoned discourse as scientists do. Epistemic knowledge is required to understand the distinction between observations, facts, hypotheses, models and theories, but also to understand why certain procedures, such as experiments, are central to establishing knowledge in science.

Content areas subscales

Knowledge can also be classified according to the major scientific fields to which it pertains. Fifteen-year-old students are expected to understand major explanatory ideas and theories from the fields of physics, chemistry, biology, earth and space sciences, and how they apply in contexts where the elements of knowledge are interdependent or interdisciplinary.

Items used in the assessment are classified into three content areas: *physical systems*, *living systems*, and *earth and space systems*. Examples of knowledge that 15-year-olds are expected to have acquired include an understanding of the particle model of matter (physical systems), the theory of evolution by natural selection (living systems), and the history and scale of the universe (earth and space systems). About 61 of all the science-related items in PISA 2015 relate to physical systems, 74 to living systems, and the remaining 49 to earth and space systems.

Quality assurance

The results of adjudication and subsequent further examinations showed that the PISA technical standards were met in all countries and economies that distributed the PISA 2015 teacher questionnaire, except Malaysia.

In Malaysia, the PISA assessment was conducted in accordance with the operational standards and guidelines of the OECD. However, the weighted response rate among the initially sampled Malaysian schools (51%) falls well short of the standard PISA response rate of 85%. Therefore, the results may not be comparable to those of other countries or to results for Malaysia from previous years.

The sample for Macao (China) represents a census of all 15-year old students and, as a consequence, of all schools attended by 15-year olds. Allowing for very limited non-response among the sampled students, the sampling error for this partner economy is minimal since the entire population of eligible students and schools are part of the surveyed sample. This, in turn, results in very small standard errors on the statistics presented in this paper.

Missing data

PISA 2015, like any other survey, suffers from missing data resulting from non-response to particular questions in the student, school or teacher questionnaires. Missing values were discarded from the statistical computations. This is referred to in statistical literature as casewise deletion. Note that all analyses involving the same variables (i.e. regressions, cross-tabulations and descriptive statistics) are based on the same samples.

Quartiles

In order to simplify the analyses, students and schools were classified into quartiles on some of the PISA standardised indices and other continuous variables. Quartiles are the three values (of a variable) that divide the data into four equal groups with each containing 25% of the observations.

For instance, in order to classify students on a scale of school socio-economic affluence, the student index of economic, social and cultural status (ESCS) is averaged at the school level and then the quartiles of the school average ESCS are computed. By doing so, students are divided into four groups depending on the socio-economic status of their school. The bottom quartile contains the 25% of students attending the most disadvantaged schools, while the top quartile contains the quarter attending the most advantaged ones.

Odds ratios

The odds ratio is a measure of the relative likelihood of a particular outcome across two groups. The odds ratio for observing the outcome when an antecedent is present is simply

$$OR = \frac{(p_{11}/p_{12})}{(p_{21}/p_{22})}$$

where p_{11}/p_{12} represents the "odds" of observing the outcome when the antecedent is present, and p_{21}/p_{22} represents the "odds" of observing the outcome when the antecedent is not present. Logistic regression can be used to estimate the log ratio: the exponentiated logit coefficient for a binary variable is equivalent to the odds ratio. A "generalised" odds ratio, after accounting for other differences across groups, can be estimated by introducing control variables in the logistic regression.

Standard errors and significance tests

The statistics in this report represent estimates based on samples of students, rather than values that could be calculated if every student in every country had answered every question. Consequently, it is important to measure the degree of uncertainty of the estimates. In PISA, each estimate has an associated degree of uncertainty, which is expressed through a standard error. The use of confidence intervals provides a way to make inferences about the population 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.

In the tables and charts used in this report, differences are labelled as statistically significant if the probability of reporting a difference when there is actually no such difference in corresponding population values is lower than 5%. Similarly, the risk of reporting a correlation as significant if there is, in fact, no correlation between two measures, is contained at 5%.

Differences between subgroup means

Differences between groups of students were tested for statistical significance. The definitions of the subgroups can, in general, be found in the charts and the text accompanying the analysis. All differences marked in bold in the tables presented in Annex B of this report are statistically significant at the 95% level.

Change in the performance per unit of an index

For many tables, the difference in student performance per unit of an index was calculated. Numbers in bold indicate that the differences are statistically significantly different from zero at the 95% confidence level.

Causation vs. correlation

The strength of PISA lies in its coverage of student, school and teacher contexts and in the comparative nature of the data. However, in non-experimental, cross-sectional data such as those gathered through PISA, even sophisticated statistical methods cannot identify cause-and-effect relationships between teaching practices and student outcomes. In particular, the inability to identify causal effects arises from the non-random selection of students and teachers into schools and from reverse causation.

Since teachers and students are not randomly assigned to schools, relationships between teaching practices and student outcomes could be confounded by factors not accounted for in the analyses, such as teachers' or students' attitudes and behaviours. The second challenge arises from reverse causation if student characteristics, such as being a low performer, affect the type of teaching practice the teacher adopts (e.g. levels of teacher support). Both challenges would lead to results that should not be interpreted as causal.

In order to limit selection bias and reverse causation, a school fixed-effects approach is adopted. This approach allows the researcher to account for all school-related observed and unobserved characteristics. In addition to school observed and unobserved characteristics, the models account for students' socio-economic status, gender, grade and number of science subjects studied.

Annex B

Student-level indices

Immigrant background

The PISA database contains three country-specific variables relating to the country of birth of the students, their mother and their father (COBN_S, COBN_M, and COBN_F). The items ST019Q01TA, ST019Q01TB and ST019Q01TC were recoded into the following categories: (1) country of birth is the same as country of assessment and (2) other. The index of immigrant background (IMMIG) was calculated from these variables with the following categories: (1) non-immigrant students (those students who had at least one parent born in the country in which they sat the assessment); (2) second-generation immigrant students (those born in the country of assessment but whose parent[s] were born in another country); and (3) first-generation immigrant students (those students born outside the country of assessment and whose parents were also born in another country). Students with missing responses for either themselves or for both parents were assigned missing values for this variable.

Based on this derived variable, a binary indicator was computed for schools in which more than 30% of students have a first- or second-generation immigrant background.

Language spoken at home

Students indicated what language they usually speak at home (ST022), and the database includes a derived variable (LANGN) containing a country-specific code for each language. In addition, an internationally comparable variable (ST022Q01TA) was derived from this information with the following categories: (1) language spoken at home is the same as the language of assessment for that student and (2) language spoken at home is another language.

Based on this derived variable, a binary indicator was computed for schools in which more than 30% of students speak a language different from that of the assessment.

Science-related career expectations

In PISA 2015, students were asked "what kind of job they expect to have when they are about 30 years old" (ST114). Answers to this open-ended question were coded to four-digit ISCO (International Standard Classification of Occupations) codes (ILO, 2007), in variable OCOD3. This variable was used to derive the index of science-related career expectations.

Science-related career expectations are defined as those career expectations whose realisation requires further engagement with the study of science beyond compulsory education, typically in formal tertiary education settings. The classification of careers into science-related and non-science-related is based on the four-digit ISCO-08 classification of occupations.

Only professionals (major ISCO group 2) and technicians/associate professionals (major ISCO group 3) were considered to fit the definition of science-related career expectations. In a broad sense, several managerial occupations (major ISCO group 1) are clearly science-related: these include research and development managers, hospital managers, construction managers, and other occupations classified under production and specialised services

managers (sub-major group 13). However, it was considered that when science-related experience and training is an important requirement of a managerial occupation, these are not entry-level jobs, and 15-year-old students with science-related career expectations would not expect to be in such a position by the age of 30.

Several skilled agriculture, forestry and fishery workers (major ISCO group 6) could also be considered to work in science-related occupations. The United States O*NET Online (2016) classification of science, technology, engineering and mathematics (STEM) occupations indeed include these occupations. These, however, do not typically require formal science-related training or study after compulsory education. On these grounds, only major occupation groups that require ISCO skill levels 3 and 4 were included among science-related occupational expectations.

Among professionals and technicians/associate professionals, the boundary between science-related and non-science-related occupations is sometimes blurred, and different classifications draw different lines.

The classification used in this paper includes four groups of jobs:

1. *Science and engineering professionals*: All science and engineering professionals (sub-major group 21), except product and garment designers (2163), graphic and multimedia designers (2166).
2. *Health professionals*: All health professionals in sub-major group 22 (e.g. doctors, nurses, veterinarians), with the exception of traditional and complementary medicine professionals (minor group 223).
3. *ICT professionals*: All information and communications technology professionals (sub-major group 25).
4. *Science technicians and associate professionals*, including:
   - physical and engineering science technicians (minor group 311)
   - life science technicians and related associate professionals (minor group 314)
   - air traffic safety electronic technicians (3155)
   - medical and pharmaceutical technicians (minor group 321), except medical and dental prosthetic technicians (3214)
   - telecommunications engineering technicians (3522).

Interest in science

The index of broad interest in science topics (INTBRSCI) was constructed using students' responses to a new question developed for PISA 2015 (ST095). Students reported on a five-point Likert scale (with the responses "not interested", "hardly interested", "interested", "highly interested", and "I don't know what this is") their interest in the following topics: biosphere (e.g. ecosystem services, sustainability); motion and forces (e.g. velocity, friction, magnetic and gravitational forces); energy and its transformation (e.g. conservation, chemical reactions); the Universe and its history; and how science can help prevent disease. The last response category ("I don't know what this is") was recoded as a missing for the purpose of deriving the index INTBRSCI. Higher values on the index reflect greater levels of agreement with these statements.

Enjoyment of science

The index of enjoyment of science (JOYSCIE) was constructed based on a trend question (ST094) from PISA 2006 (ID in 2006: ST16), asking students, on a four-point Likert scale (with the responses "strongly agree", "agree", "disagree", and "strongly disagree"), about their agreement with the following statements: "I generally have fun when I am learning <broad science> topics"; "I like reading about <broad science>"; "I am happy working on <broad science> topics"; "I enjoy acquiring new knowledge in <broad science>"; and "I am interested in learning about <broad science>". The derived variable JOYSCIE was equated to the corresponding scale in the PISA 2006 database, thus allowing for a trend comparison between PISA 2006 and PISA 2015. Higher values on the index reflect greater levels of agreement with these statements.

Science self-efficacy

The index of science self-efficacy (SCIEEFF) was constructed based on a trend question (ST129) that was taken from PISA 2006 (ID in 2006: ST17). Students were asked, using a four-point answering scale (with the responses "I could do this easily"; "I could do this with a bit of effort"; "I would struggle to do this on my own"; and "I couldn't do this") to rate how they would perform in the following science tasks: recognise the science question that underlies a newspaper report on a health issue; explain why earthquakes occur more frequently in some areas than in others; describe the role of antibiotics in the treatment of disease; identify the science question associated with the disposal of garbage; predict how changes to an environment will affect the survival of certain species; interpret the scientific information provided on the labelling of food items; discuss how new evidence can lead you to change your understanding about the possibility of life on Mars; and identify the better of two explanations for the formation of acid rain. Responses were reverse-coded so that higher values of the index correspond to higher levels of science self-efficacy. The derived variable SCIEEFF was equated to the corresponding scale in the PISA 2006 database, thus allowing for a trend comparison between PISA 2006 and PISA 2015.

Disciplinary climate

The index of disciplinary climate (DISCLISCI) was constructed from students' reports on how often ("every lesson", "most lessons", "some lessons", "never or hardly ever") the following happened in their science lessons (ST097): Students don't listen to what the teacher say; There is noise and disorder; The teacher has to wait a long time for students to quiet down; Students cannot work well; Students don't start working for a long time after the lesson begins.

Epistemic beliefs about science

The index of epistemic beliefs about science (EPIST) was constructed using students' responses to a new question developed for PISA 2015 about students' views on scientific approaches (ST131). Students reported, on a four-point Likert scale (with the responses "strongly disagree", "disagree", "agree", and "strongly agree"), their agreement with the following statements: A good way to know if something is true is to do an experiment; Ideas in <broad science> sometimes change; Good answers are based on evidence from many different experiments; It is good to try experiments more than once to make sure of your findings; Sometimes <broad science> scientists change their minds about what is true in science; and The ideas in <broad science> science books sometimes change. Higher levels on the index correspond to greater levels of agreement with these statements.

PISA index of economic social and cultural status (ESCS)

The PISA index of economic, social and cultural status (ESCS) was derived, as in previous cycles, from three variables related to family background: parents' highest level of education (PARED), parents' highest occupation status (HISEI), and home possessions (HOMEPOS), including books in the home. PARED and HISEI are simple indices, and HOMEPOS is a proxy measure for family wealth.

For the purpose of computing the PISA index of economic, social and cultural status (ESCS), values for students with missing PARED, HISEI or HOMEPOS were imputed with predicted values plus a random component based on a regression on the other two variables. If there were missing data on more than one of the three variables, ESCS was not computed and a missing value was assigned for ESCS.

The PISA index of economic, social and cultural status was derived from a principal component analysis of standardised variables (each variable has an OECD mean of zero and a standard deviation of one), taking the factor scores for the first principal component as measures of the PISA index of economic, social and cultural status. All countries and economies (both OECD and partner countries/economies) contributed equally to the principal component analysis, while in previous cycles, the principal component analysis was based on OECD countries only. However, for the purpose of reporting, the ESCS scale was transformed, with zero being the score of an average OECD student and one being the standard deviation across equally weighted OECD countries.

Principal component analysis was also performed for each participating country or economy separately, to determine the extent to which the components of the index operate in similar ways across countries or economies.

School-level indices

School type

Schools are classified as either public or private according to whether a private entity or a public agency has the ultimate power for decision making concerning its affairs (SC013). As in previous PISA surveys, the index on school type (SCHLTYPE) has three categories, based on two questions: SC013, which asks if the school is a public or a private school, and SC016, which asks about the sources of funding. This index was calculated in 2015 and in all previous cycles.

Annex C

List of tables

All tables in Annex B together with other results are available online at:
http://www.oecd.org/pisa/sitedocument/PISA-2015-Working-Paper-Teaching-Science-Data.xlsx.