101547

PISA

The Experience of Middle-Income Countries Participating in PISA 2000-2015

Programme for International Student Assessment

OECD

WORLD BANK GROUP

PISA

The Experience of Middle-Income Countries Participating in PISA 2000-2015

Marlaine E. Lockheed
with
Tijana Prokic-Breuer
Anna Shadrova

The opinions expressed and arguments employed herein are solely those of the authors and do not necessarily reflect the official views of the OECD, its member countries, the World Bank, its Board of Executive Directors, or of the governments they represent.

This document and any 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.

Please cite this publication as:
Lockheed, M., T. Prokic-Bruer and A. Shadrova (2015), *The Experience of Middle-Income Countries Participating in PISA 2000-2015,* PISA, World Bank, Washington, D.C./OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789264246195-en

ISBN 978-92-64-24618-8 (print)
ISBN 978-92-64-24619-5 (PDF)

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

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.

The names of countries and territories used in this joint publication follow the practice of the OECD.

Photo credits:
© epicurean/iStockphoto
© Ian Lishman/Juice Images/Inmagine LTD
© Istockphoto/Henk Badenhorst
© Steve Debenport/iStockphoto

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

Foreword

Although large numbers of children in developing countries have been able to enter school over the past two decades, many young people – especially the disadvantaged, young girls, those living in rural areas and ethnic minorities – are still leaving school without the knowledge and skills they need to thrive in society and find decent livelihoods. This has led to a general consensus[1] that the 2030 Sustainable Development Goal for education should include a focus on learning. That, in turn, makes reliable metrics on the quality of learning outcomes truly essential.

Every three years, some 80 countries and economies collaborate to compare how well their school systems prepare young people for life and work. The framework for these comparisons is an international assessment of the knowledge and skills of 15-year-old students known as PISA, the Programme for International Student Assessment. PISA does not just examine whether students have learned what they were taught, but also assesses whether students can creatively and critically use what they know.

Of course, such international comparisons are never easy and they aren't perfect. But they show what is possible in education, they help governments to see themselves in comparison to the education opportunities and results delivered by other education systems, and they help governments to build effective policies and partnerships for improving learning outcomes.

But as the number of countries joining PISA kept rising, it became apparent that the design and implementation models for PISA needed to evolve to successfully cater for a larger and more diverse set of countries. This publication provides systematic analyses of these issues. Chapter 2 describes the pattern of country participation in PISA and other large-scale international assessments over time, as well as the broad determinants of PISA participation. Chapter 3 examines the financial, technical and contextual challenges for countries participating in PISA. Chapter 4 explores the capacity-building outcomes for countries participating in PISA. Chapter 5 examines how PISA results have informed international and national policy discussions. Finally, Chapter 6 explores PISA data from middle-income countries, particularly with respect to the quality and equity of education in these contexts.

In response to these challenges, the OECD and a number of partners have launched the PISA for Development initiative that aims to extend PISA to support evidence-based policy making in emerging and developing economies and to help them monitor progress towards the 2030 Sustainable Development Goals.

A first aim is to further develop and differentiate the PISA contextual questionnaires and data-collection instruments in order to better capture the diverse contexts in which students learn, teachers teach and schools and school systems operate in emerging and developing countries. This will allow for a deeper understanding of how certain factors – such as the socio-economic background of students or the learning environment in classrooms – are associated with learning outcomes in different contexts.

A second aim is to make the PISA assessment instruments more sensitive to a wider range of performance levels. While there are undoubtedly high performers in all countries, a number of 15-year-old students in developing countries can be expected to perform at lower levels of proficiency. The enhanced assessment instruments will better capture performance differences among these students and cater for more diverse cultural contexts, while maintaining the comparability of a country's results on the global PISA scales.

Perhaps the biggest challenge for PISA is to ensure that those children who still do not have access to schooling remain no longer beyond the reach of the metrics used to evaluate the success of education systems. Though much progress has been made in increasing access to education around the world, over 70 million children of lower-secondary-school age remain out of school. An innovative out-of-school assessment will allow future PISA assessments to provide data on learning outcomes for those children too.

1. See UNESCO and UNICEF (2013), *Making Education a Priority in the Post-2015 Development Agenda*, www.worldwewant2015.org/node/389575.

Together, these efforts will enable more countries to use PISA to set national learning targets, monitor progress towards them, and analyse the factors that affect student outcomes, particularly among poor and marginalised populations.

PISA is also putting a new partnership model into practice: a group of countries from the developing world is leading the development of these new methodologies and collaborating on finding ways for making the most effective use of the results for the design and implementation of policies. They will develop the enhanced survey instruments and methodologies and undertake field trials and surveys to test them. The results of their surveys will provide local policy makers with new evidence to diagnose shortcomings and inform new policies. In addition, these countries will benefit from opportunities for peer-to-peer exchanges with other members of the PISA global community.

Within each partner country, bilateral and multilateral development co-operation agencies will provide both financial and technical support for the initiative, including assistance in developing the survey instruments and methodologies. The results of the survey can be used to inform their own policies and practices.

Technical partnerships and networks of experts in such fields as learning assessments or out-of-school youth will be established to build on best practices. In addition to offering insights into learning metrics at global and national levels, these partnerships will also work to address the lack of data on out-of-school children.

The OECD remains committed to maintain and develop PISA as a global yardstick for educational success. It will continue to contribute its expertise and its platforms to support collaboration for developing and conducting PISA surveys and to assist policy-makers and practitioners throughout the world to use them most productively.

Andreas Schleicher
Director, Directorate for Education and Skills,
OECD

Claudia Costin
Senior Director for the Education Global Practice,
World Bank

Acknowledgements

This report was produced with the support of the World Bank through its Russia Education Aid for Development (READ) Trust Fund programme, as part of its contribution to the PISA for Development project. PISA for Development is an initiative of the OECD and development partners that aims to identify how PISA can best support evidence-based policy making in emerging and developing economies and contribute to the UN-led definition of global learning goals for the post-2015 agenda. In addition, the PISA for Development project will help to build country capacity in assessment, analysis and use of results for monitoring and improvement among participating countries.

This report was authored by Marlaine E. Lockheed, former World Bank official, with researchers Tijana Prokic-Breuer and Anna Shadrova as part of the PISA for Development project.

The authors wish to thank Michael Ward, OECD, and Marguerite Clarke, World Bank, for suggesting this project. Without their support and guidance throughout it would not have been possible. Pablo Zoido, OECD, also provided valuable input and guidance throughout. In addition, the World Bank's Education Practice Managers provided much appreciated guidance early on in terms of key questions to be addressed by the study and the choice of case study countries. The report also benefitted from the critical comments and suggestions of peer reviewers, including the OECD's Jenny Bradshaw and Sophie Vayssettes together with other members of the OECD's PISA team, members of the PISA Strategic Development Group (SDG), and members of the PISA for Development International Advisory Group. Editorial and administrative services were provided by Jennifer Cannon, Hélène Guillou and Kelly Makowiecki, OECD, and the layout work was done by Fung Kwan Tam.

The authors also wish to thank all the PISA 2015 National Project Manager teams from middle-income countries who gave their time to meet with us and answer numerous follow-up emails; the World Bank staff, former staff and Millennium Challenge Corporation staff who provided information regarding PISA in the case-study countries; and the contractors for PISA 2015 and PISA for Development who allowed us to sit in on their meetings and participated in interviews. Thanks also to the UNESCO Global Monitoring Report 2015 Director, Aaron Benavot, for providing the global data on national assessments used in Chapter 2 of this report. Finally, we wish to acknowledge the long and fruitful conversations with the late Eugene Owen, the first Chair of the PISA Governing Board, for his wisdom regarding PISA's initial years.

Table of contents

TABLE OF CONTENTS

BOXES

FIGURES

TABLES

Abbreviations and acronyms

ACER	Australian Council for Education Research
APA	American Psychological Association
CEQE	Center of Education Quality Evaluation
CNA	Capacity Needs Analyses
CONFEMEN	Conférence des Ministres de l'Education des Pays Ayant le Francais en Partage
ERCE	Estudio Regional Comparativo y Explicativo (Regional Comparative and Explanatory Study); also shown as First ERCE, (or PERCE), Second ERCE (SERCE), and Third ERCE (TERCE)
ERDD	Educational Research and Development Department
ERfKE	Education Reform for the Knowledge Economy
ETS	Educational Testing Service
EU	European Union
FYROM	Former Yugoslav Republic of Macedonia
GDN	Global Development Network
GDP	Gross domestic product
GRADE	Groupo de Analisis para el Desarrollo (Analysis for Development Group)
IAEP	International Assessment of Education Progress
ICCS	International Computer Competence Study
ICFES	Colombian Institute for Educational Evaluation
ICT	Information and Communications Technology
IEA	International Association for the Evaluation of Educational Achievement
IERI	IEA-ETS Research Institute
IIEP	International Institute for Educational Planning
ILSA	International large-scale assessments
IMF	International Monetary Fund
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais (National Institute of Educational Studies and Research)
IRLS	International Reading Literacy Study
ISCED	International Standard Classification of Education
LLECE	Laboratorio Latinamericano de Evaluacion de la Calidad de la Education (Latin American Laboratory for the Evaluation of Educational Quality)
MCC	Millennium Challenge Corporation
NAEC	National Assessment and Examinations Center
NCHRD	National Center for Human Resources Development
NLSA	National large-scale assessment
NPM	National Project Manager
NRC	National Research Co-ordinator
PASEC	Programme d'Analyse des Systèmes Educatif de CONFEMEN
PIAAC	Programme for the International Assessment of Adult Competencies
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PREAL	Partnership for Educational Revitalization in the Americas
RCT	Randomised Control Trial
RLSA	Regional large-scale assessment
SABER	Systems Approach for Better Education Results
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SAEB	Sistema de Avaliação da Educação Básica (Basic Education Evaluation System)
SIMCE	Sistema de Medicion de la Calidad de la Education (Education Quality Measurement System)
TALIS	Teaching and Learning International Study
TIMSS	Trends in International Mathematics and Science Study
UNDP	United National Development Programme
UNESCO	United Nations Education, Science and Culture Organization
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
VNIES	Vietnam National Institute for Education Sciences

PISA FOR DEVELOPMENT

Executive Summary

The Education 2030 agenda that is set within the framework of the Sustainable Development Goals adopted by the UN in September 2015 emphasises the quality, equity and measurement of learning outcomes for young children through to working adults. This report serves to inform Education 2030 discussions through an analysis of the OECD's Programme for International Student Assessment (PISA) as it has been experienced in middle-income countries.

PISA is a triennial internationally comparable assessment of 15-year-olds in school. It measures student performance in reading, mathematics and science, and collects contextual data through school and student questionnaires. A random sample of at least 5 000 students are tested in each country.

The OECD is seeking to enhance PISA to make it more relevant to a wider range of countries. The PISA for Development initiative has been designed to enable greater PISA participation by low- and middle-income countries. This report presents evidence from the experience of the 4 low-income and 40 middle-income countries, including 3 OECD member countries that have already participated in one or more cycles of PISA.

The report presents evidence that:

- **More countries are participating in all forms of assessments,** including international large-scale assessments such as PISA. Approximately two-thirds of all countries with populations greater than 30 000 have participated in one or more international or regional large-scale assessments.
- **Participation in PISA, specifically, has also increased, but wealthier countries have participated more often than less wealthy countries**. All OECD countries participate in PISA, many upper-middle-income countries participate in PISA, but very few low-income and lower-middle-income countries have participated in PISA.
- **Middle-income countries, particularly lower-middle-income countries, encounter financial, technical and contextual challenges when participating in PISA**. Financial challenges are the most pressing, with the main financial challenge being adequate and sustained financing for the units or agencies responsible for the assessment. Technical challenges involve some specific PISA tasks, particularly for translation, sampling, survey administration and coding of responses. Contextual challenges involve the political, regulatory and cultural environment of the countries, which can affect ease of assessment implementation and use.
- **PISA participation has built on prior assessment experience.** To date, most countries that have participated in PISA have had significant prior experience with other international large-scale assessments or with national assessments, allowing PISA implementation to build on this experience and related expertise.
- **PISA assumes that participating countries have substantial assessment capacity.** Most capacity development has focused on specific technical quality aspects of the assessment and is closely embedded in twice-yearly meetings of the country participants.

- **PISA results have informed some education policy in middle-income countries.** Media coverage of PISA results is less extensive in middle-income countries than in OECD countries, but PISA results are used by donors in country dialogue with middle-income countries. The main reported effects have been related to agenda setting and the expression of intentions to reform curricula, standards and teacher professional development.
- **PISA results underscore differences between high-income and middle-income countries in education quality and equity, and their correlates**. Analyses of PISA 2012 for the 18 middle-income countries participating in this cycle provide evidence regarding the overall quality of and equity within education systems, as well as evidence regarding systemic differences among the countries.

The report draws five conclusions:

1. PISA will need substantial enhancements to achieve the OECD's aim of making it more relevant to a wider range of countries.
2. PISA participation challenges are greater for lower-middle-income countries than for upper-middle-income and high-income countries.
3. Much remains to be learned about the impact of PISA participation on assessment capacity development in middle-income countries.
4. Private dialogue between government officials and development partners regarding the results of PISA has been useful in helping shape investment priorities and policies in middle-income countries.
5. PISA provides valuable information about the level and distribution of quality and equity within a country's education system.

The report also makes six recommendations for countries, the OECD and other development partners:

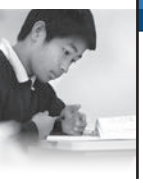
1. Countries should continue to work towards the establishment of an assessment culture that seeks to learn from international large-scale assessments, such as PISA and others, by fully funding and sustaining the costs associated with participation in international large-scale assessments and encouraging greater discussion and reporting of the results.
2. The OECD should acknowledge the differences between the countries currently participating in PISA and the low- and middle-income countries that are not currently participating by revising cognitive instruments, background questionnaires and school questionnaires.
3. The OECD should help reduce the technical challenges to PISA participation for lower-middle-income countries by extending and enhancing the training related to PISA implementation.
4. The OECD should evaluate the impact of PISA participation on assessment capacity development in middle-income countries by sponsoring a rigorous impact evaluation.
5. The World Bank and other development partners should help reduce the financial challenges to PISA participation for lower-middle-income countries by supporting countries through grants and loans.
6. The World Bank and other development partners should sponsor fellowships and grants for research and development that could lead to analyses of PISA data that is more attuned to the policy needs of individual participating countries.

PISA FOR DEVELOPMENT

1

Overview: Lessons from international and regional educational assessments

This chapter provides the rationale for undertaking this study on the experiences of middle-income countries in relation to the Programme for International Student Assessment (PISA), and an overview of its analyses, conclusions and recommendations. It links PISA with the PISA for Development initiative and presents the questions the report seeks to answer. It also provides a summary of the findings from the review of evidence and the analyses conducted for the study. It draws conclusions and makes recommendations for the OECD, the World Bank and other development partners and for low- and middle-income countries.



RATIONALE FOR THIS REPORT

The Education 2030 agenda[1] that is set within the framework of the Sustainable Development Goals (SDGs) emphasises the quality, equity and measurement of learning outcomes for young children through to working adults. This report serves to inform Education 2030 discussions through an analysis of how the OECD's PISA has been experienced in middle-income countries.

PISA is a triennial internationally comparable assessment of 15-year-olds in school. It measures student performance in reading, mathematics and science, and collects contextual data through school and student questionnaires. A random sample of at least 5 000 students are tested in each country.

The OECD is seeking to enhance PISA to make it more relevant to a wider range of countries. The PISA for Development initiative (Box 1.1) has been designed to enable greater PISA participation by low- and middle-income countries.

Box 1.1 PISA for Development: Making PISA more relevant to low- and middle-income countries

In 2014, the OECD and a number of partners launched the PISA for Development initiative that aims to identify how PISA can best support evidence-based policy making in low- and middle-income economies and contribute to the UN-led definition of global learning goals for Education 2030. These objectives will be achieved in three main ways:

- Developing contextual questionnaires and data collection instruments that better capture diverse situations in emerging and developing countries.
- Adjusting the PISA test instruments so that they are sensitive to a wider range of performance levels.
- Establishing methods and approaches to include out-of-school students in the PISA assessment.

The project will also contribute to capacity building for managing large-scale student learning assessment in the participating countries: Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia. A key feature of the project is peer-to-peer learning, which will be achieved through bringing together the middle-income countries already participating in PISA with the seven PISA for Development countries.

PISA for Development is guided by an International Advisory Group composed of representatives from development and technical partners supporting the project, and the participating countries outlined above. In addition, technical experts meet regularly to develop survey instruments and methodologies.

The OECD and its partners expect these efforts will enable more countries to use PISA to set national learning targets, monitor progress towards targets, and analyse the factors that affect student outcomes, particularly among poor and marginalised populations. Countries will also have greater institutional capacity to help track Education 2030 global targets and indicators that are focused on access plus learning.

Source: OECD PISA for Development web-site at: www.oecd.org/pisa/aboutpisa/pisafordevelopment.htm.

This report complements and helps shape the PISA for Development project through presenting evidence derived from the experience of the 44 low-income and middle-income countries, including 3 OECD member countries (Chile, Mexico and Turkey) that have already participated in one or more cycles of PISA.[2]

This report also highlights the experience of a few carefully selected countries from the middle-income participants. It examines the experiences of these countries with respect to the challenges they have faced in participating in PISA, the use they have made of PISA results, the impact of their participation on their education policies and education systems, and their capacity to manage large-scale student assessments for system improvements and target-setting. The report also identifies areas where some countries may not have exploited their investment in PISA to its full potential and where participation in PISA may not have led to the outcomes or processes the country expected. It makes suggestions for how all countries, not only those in the developed world, can participate successfully in international large-scale assessments of learning outcomes. It is also intended to be a useful policy tool for facilitating discussions with countries about participating in international large-scale assessments.

OVERVIEW: THE QUESTIONS THE REPORT SEEKS TO ANSWER

This report addresses the following six questions:

1. Which countries/economies have participated in PISA and other international large-scale assessments?
2. Why do countries/economies participate in PISA?
3. What have been the challenges facing middle-income countries participating in PISA?
4. What have been the capacity building outcomes for middle-income countries participating in PISA?
5. How have PISA results informed education policy discussions and affected education policy in middle-income countries?
6. What do PISA data tell us about education in middle-income countries?

The report uses a variety of methodologies to address these questions, including: interviews, surveys, literature reviews, reviews of donor documents, analysis of online archives of major news media in selected middle-income countries, and secondary analysis of the PISA data sets. The empirical evidence gathered for this study suggests some answers to the six questions and are summarised below.

FINDINGS FROM THE REVIEW OF EVIDENCE AND ANALYSIS

More countries are participating in international large-scale assessments

Country-level participation in international large-scale assessments has grown steadily in recent decades and approximately two-thirds of all countries with populations greater than 30 000 have participated in one or more international or regional large-scale assessment. Some low- and middle-income countries began to participate in international assessments in the early 1970s and their participation has grown considerably since then, reaching a total of 65 low- and middle-income countries participating in various international or regional assessments in 2015. PISA has also grown significantly since its first cycle in 2000 and reached about 70 countries in the 2015 cycle, 26 of which were middle-income countries as of 2013. Countries in North America and Europe/Central Asia have the highest PISA participation rates, whereas countries in South Asia and sub-Saharan Africa have the lowest.

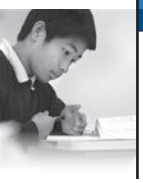
Countries participate in PISA for many reasons, but money is key

Countries decide to participate in PISA for a range of reasons. In general, wealthier countries have participated more often than less wealthy countries, and all OECD countries have participated. Most countries in Europe and North America have participated, while many countries in South Asia, Latin America, and Sub-Saharan Africa have not. Countries that have prior experience with national assessments and with other international large-scale assessments, in particular the International Association for the Evaluation of Educational Achievement's (IEA) Trends in International Mathematics and Science Study (TIMSS), have been much more likely to participate than countries lacking that experience. Experience with a regional assessment is unrelated to participation in PISA. Low-income countries[3] generally have not participated. Lower-middle-income countries have generally participated when participation has been encouraged and financed by donors. Many upper-middle-income and high-income countries have participated following long and steady experience with related activities – such as national assessments and examinations – that have established a culture of assessment with the necessary technical and administrative capacity for conducting and learning from international large-scale assessments. The evidence shows mixed experiences across countries, which suggests that context matters and that the reasons for participation in one country may be very different from the reasons in another.

Middle-income countries encounter financial, technical and contextual challenges

This report identifies three types of challenges that confront lower-middle- and upper-middle-income countries participating in PISA and other international large-scale assessments: financial challenges, technical challenges, and contextual challenges.

Financial challenges are the most pressing. The main financial challenge is adequate and sustained financing for the units or agencies responsible for PISA. Many lower-middle-income countries that have participated in PISA have benefitted from donor support for their international participation fees and other direct costs of the assessment. A second financial challenge is the modest level of information available about the direct costs of participation, with some countries significantly underestimating the level of effort and associated costs required for the successful completion of PISA. A third financial challenge relates to the absence of financial resources and support for analysing data and preparing a national report. Support for estimating the actual costs of conducting the assessment, inclusion of these costs in an annual budget, and donor support to lower-income countries could help mitigate these challenges.



The main technical challenges are largely operational: *i)* translation into multiple languages from source documents provided in English and French (also regional differences in some languages such as Spanish, French or Arabic); *ii)* sampling (particularly in large federal or decentralised systems); *iii)* survey administration; and *iv)* coding of open-ended questions. In some countries, regulations regarding contracting the part-time staff needed to implement a large-scale assessment have created challenges. Support for assessing the capacity of newly participating countries relative to the PISA standards could help mitigate this challenge.

The main contextual challenges are the political, regulatory and cultural environment of the countries. These challenges are often broad and difficult to address as they vary from country to country. Nevertheless, it may be necessary to analyse these challenges as they relate to specific countries and to make adjustments in certain PISA processes at the country level, while adhering to quality standards. In particular, additional effort may be needed for outreach to stakeholders so that the entire education system – including the central ministry, social partners, teachers, parents and students – has a better understanding of the assessment, what it shows, and how it can help guide quality and equity improvements.

PISA participation has built on prior assessment experience

To date, most countries that have participated in PISA have had significant prior experience with other international large-scale assessments or with national assessments, allowing PISA implementation to build on this experience. The capacity development activities of previous PISA cycles have been limited to specific technical quality aspects of the assessment and are closely embedded in the twice-yearly meetings of the National Project Managers. Training is designed to strengthen existing capacities with respect to operational tasks, such as sampling, data collection, data entry, professional scoring of constructed-response items and translation. In general, training has been effective. It was not designed to enhance the enabling environment for assessment, although in many cases this can be a side effect when PISA is carried out by national assessment agencies. Extending and deepening this training may be needed to enhance the assessment capacity in some low- and middle-income countries.

PISA results have informed some education policy in middle-income countries

Policy formulation begins with agenda setting, often through the public discussion of policy issues in the media. Specific policy reforms in low- and middle-income countries are also discussed in the context of private, country-level dialogue with donors. Evidence indicates that high-income OECD member countries have responded to the publication of PISA results by seeking to learn from the experience of other countries and by reflecting on education policy, often very publicly. Available evidence suggests that this is less the case in middle-income countries, where public dialogue about education policy, as observed in selected public media from eight middle-income countries, has little relationship with the publication of PISA results.

However, PISA and other international large-scale assessments have informed the policy dialogue between countries and donors, as indicated by a review of World Bank projects. International large-scale assessments have provided empirical evidence of the need for policy reforms to improve the quality and equity of education outcomes in numerous countries, and have occasionally been used as key results indicators for projects and programmes supported by donors. Most empirical evidence of PISA's direct impact on policy reforms focuses on high-income countries. Some evidence for middle-income countries suggests effects on curricula, performance standards and assessments.

PISA results underscore differences between high-income and middle-income countries in education quality and equity

Analyses of PISA 2012 for all participating low- and middle-income countries provide evidence regarding the overall quality of and equity within education systems, as well as evidence regarding systemic differences among the countries. These analyses also demonstrate differences between high- and middle-income countries and among countries and regions regarding the factors associated with higher performance. For example, the socio-economic background of students explains more between-school differences in middle-income countries than in OECD countries, on average. These analyses also indicate, however, that existing PISA student background and school questionnaires may be more relevant for high-income countries than for low-income or lower-middle-income countries, since – in particular – the school questionnaire provides little information regarding the types of resources that often explain between- or within-school differences in student performance in developing countries. Both of these questionnaires may need to be modified to better reflect the home environments and school contexts common in low-income and lower-middle-income countries.

CONCLUSIONS

The report reaches five conclusions:

1. **PISA will need substantial enhancements to achieve the OECD's aim of making it more relevant to a wider range of countries.** Virtually all high-income countries and many upper-middle-income countries already participate in PISA. Participation in PISA is associated with national wealth, geographical location in Europe/Central Asia and North America, and prior experience with national or international large-scale assessments. Non-participating low- and middle-income countries may lack some or all of these factors, and therefore they may find participation more challenging.
2. **PISA participation challenges are greater for lower-middle-income countries than for upper-middle-income countries and high-income countries**. These challenges include financial, technical and contextual challenges in conducting the assessment, which are typically not challenges to higher-income countries. The two main financial challenges are adequate and sustained financing for the units or agencies responsible for PISA, and insufficient information regarding the incremental costs associated with participation.
3. **Much remains to be learned about the impact of PISA participation on assessment capacity development in middle-income countries.** PISA implementation has benefitted from existing assessment capacity in most participating countries, developed through participation in other assessments. To date, there has been no rigorous evaluation of the impact of PISA participation on assessment capacity development, particularly for middle-income countries. Some middle-income countries, however, report that participation has built some operational capacity. To answer the question of capacity development, a sound evaluation is needed.
4. **Private dialogue between government officials and development partners regarding the results of PISA has been useful in helping shape investment priorities and policies in middle-income countries**. The role of PISA in affecting education policy may differ between middle-income countries and OECD countries, since public, published discussions about education policy are less prevalent in some middle-income countries.
5. **PISA provides valuable information about the level and distribution of quality and equity within a country's education system.** However, the existing PISA cognitive instruments and questionnaires need to be enhanced to better capture the range of performance in low- and middle-income countries, and to provide more information regarding context, institutions and inputs that could improve education quality and equity in middle-income countries.

RECOMMENDATIONS

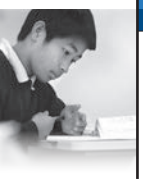
On the basis of the above findings, this report makes recommendations directed to low-income and middle-income countries, the OECD, and to the World Bank and other development partners. These recommendations are intended to: *i)* enhance the future development of PISA in ways that could help low- and middle-income countries better participate in PISA and other international large-scale assessments; *ii)* encourage financial support for greater participation in PISA and other international assessments by low- and middle-income countries; and *iii)* improve the utility of PISA participation by low- and middle-income countries.

Recommendations for low- and middle-income countries

1. Continue to work towards the establishment of an assessment culture that seeks to learn from international large-scale assessments, such as PISA and others, by:
   - Fully funding the incremental implementation costs for participation in international large-scale assessments through sustained support to implementing units or agencies.
   - Increasing outreach to stakeholders through both conventional media, new media and seminars, so that the entire education system – from the central ministry to teachers to parents – has a better understanding of the assessment, what it shows, and how it can help guide quality and equity improvements.

Recommendations for the OECD[4]

2. Acknowledge the differences between countries currently participating in PISA and the low- and middle-income countries that are not currently participating by:
   - Expanding the cognitive instruments to reflect where the majority of students in developing countries are likely to be in terms of learning levels and, in particular, to better capture performance differences at the lower end of the performance scales.



- Enhancing the questionnaires to better measure contexts, institutions and inputs that are relevant to these countries, through improving the quality of information collected on teachers and teaching, emphasising features of schools that are more important in low- and lower-middle-income countries (such as basic infrastructure and instructional materials), and enhancing the measure of students' socio-economic background.
- Reviewing the implementation process to better understand constraints in low- and middle-income countries and propose new approaches that maintain current levels of rigour and quality while making participation more manageable and attractive.
- Expanding some of the regular OECD analyses of PISA to focus on middle-income participating countries. This can be done by analysing results separately by income group and world region.

3. Help reduce the technical challenges to PISA participation by lower-middle-income countries by:
   - Providing source documents in additional languages, with verified translations.
   - Providing capacity assessment, capacity building and increased continuous remote assistance from international contractors hired by the OECD to support countries as they implement the tasks associated with the assessment.
   - Analysing the contextual challenges for each new participating country and adjusting PISA processes as needed, while maintaining standards for the reliability and validity of the assessment results.
4. Evaluate the impact of PISA participation on assessment capacity development in middle-income countries by:
   - Implementing baseline assessments of country capacity and monitoring change in capacity over time for low- and middle-income countries joining PISA 2018.

Recommendations for the World Bank and other development partners

5. Help reduce the financial challenges to PISA participation for lower-middle-income countries by:
   - Supporting the development of a document, aligned with the National Project Manager's manual, to help countries better estimate the direct costs associated with each task required for implementing the assessment.
   - Using country dialogue to encourage countries to fully fund the implementation costs of PISA and other large-scale assessments in the annual budgets of the units or agencies responsible for conducting the assessment.
   - Continuing to support the costs of participation in international large-scale assessments through the World Bank's lending operations and grants programmes and development co-operation more broadly.
6. Help support research and development that could lead to analyses of PISA data more attuned to the policy needs of the individual participating countries by:
   - Establishing fellowship and grants programmes for this purpose, building on the experience accumulated at the OECD with the Thomas J. Alexander Fellowship Programme,[5] which focuses on emerging economies.
   - Financing or providing technical training for researchers in low- and lower-middle-income countries in the use of data sets from PISA and other large-scale international assessments.
   - Financing the preparation of country-level reports based on in-depth research using PISA data.

Notes

1. The Education 2030 agenda refers to the fourth Sustainable Development Goal (SDG) – ensure inclusive and equitable quality education and promote life-long learning opportunities for all – and the ten targets that have been put in place to support this. This goal and targets were adopted by the UN General Assembly, along with 16 other SDGs and 159 other targets, at the sustainable development summit in New York in September 2015.

2. 44 countries have participated in some PISA cycle at the time they were classified as low- or middle-income; not all these countries are currently low- or middle-income countries.

3. This report was originally designed to include issues of low-income countries participating in PISA, but since only four countries have ever participated at the time they were classified as low-income countries, the report instead focuses on lower-middle-income and upper-middle-income countries.

4. Several of these recommendations were anticipated by the ongoing PISA for Development initiative.

5. The Thomas J. Alexander Fellowship Programme of the OECD seeks to develop and support improvements in education quality and equity, particularly in emerging economies, in three ways by: supporting quantitative research to provide evidence-based analysis relating to education policy and practice; strengthening the use of such analysis in policy making; and helping education leaders modify their practices to maximise its impact on student learning.

PISA FOR DEVELOPMENT

2

International large-scale assessments: Origins, growth and why countries participate in PISA

This chapter describes the origins of international large-scale assessments, presents evidence regarding the worldwide growth in such assessments, and analyses the broad determinants of participation in the Programme for International Student Assessment (PISA). It links PISA participation with countries' economic status and their experience with national and other international large-scale assessments. The empirical analyses presented in the second section of this chapter use selected time-series data for over 200 countries/economies to empirically estimate the determinants of PISA participation. The third section examines country-by-country variations in PISA participation over time. The fourth section draws on interviews with various policy actors to explore the PISA experience of selected low- and middle-income case study countries. A final section draws conclusions about what factors contribute to a country's decision to participate in PISA and implications for the expansion of PISA participation.

ORIGINS OF INTERNATIONAL LARGE-SCALE ASSESSMENTS

Fifty-five years ago, a group of scholars from Western Europe and North America met at the United Nations Education, Science and Culture Organization (UNESCO) Institute for Education in Hamburg, Germany, to explore the feasibility of an international study of educational achievement. Twelve countries, nine high-income and three upper-middle-income, participated in what was later referred to as the International Association for the Evaluation of Educational Achievement (IEA) Pilot Study.[1] These countries were: Belgium (French Community), England, Germany, Finland, France, Israel, Poland, Scotland, Sweden, Switzerland, the United States and Yugoslavia (Husen, 1967; Keeves, 2011). As Keeves noted: "The Pilot Study established that a cross-national investigation was feasible in spite of the problems of translation and administration." This conclusion has underpinned the remarkable rise of international large-scale assessments over the past five decades, despite the gradual inclusion of countries that dramatically differ in terms of economy and culture from those initial twelve countries.

Lower income countries in international large-scale assessments, 1960s-90s

Low- and middle-income countries[2] began to participate in international assessments in the early 1970s. The IEA Pilot Study was followed in 1964 by IEA's First International Mathematics Study in which twelve high-income countries participated: Australia, Belgium (both Flemish and French Communities), England, Germany, Finland, France, Israel, Japan, Netherlands, Scotland, Sweden and the United States (Husen, 1967). In 1970-71, up to 19 countries participated in IEA's Six-Subject Study, covering the domains of science, reading, literature, French as a foreign language, English as a foreign language and civic education. For the first time, participating countries included six low- or middle-income countries: Chile, Hungary, India, Iran, Romania and Thailand.

A decade later, IEA's Second Mathematics Study was initiated, with four low- or middle-income countries participating: Hungary, Nigeria, Swaziland and Thailand (Robitaille and Garden, 1989). In 1983-85, IEA's Second Science Study was carried out, with eight low- or middle-income countries participating: three provinces of the People's Republic of China (hereafter 'China'), Ghana, Hungary, Nigeria, Papua New Guinea, Philippines, Thailand and Zimbabwe (Postlethwaite and Wiley, 1992). In 1988, Educational Testing Service carried out its International Assessment of Educational Progress (IAEP) in 6 countries, and repeated the exercise in 1991 in 20 countries, including 5 low- or middle-income countries: Brazil, China, Hungary, Jordan and cities in Mozambique (Lapointe, Askew and Mead, 1992; Lapointe, Mead, and Askew, 1992; Lapointe, Mead, and Phillips, 1989). In 1990-91, IEA's Reading Literacy Study was carried out, with 8 low- or middle-income countries participating: Botswana, Hungary, Indonesia, Nigeria, Philippines, Thailand, Venezuela and Zimbabwe (Elley, 1992). Thus, as of a quarter-century ago, 19 low- and middle-income countries from all world regions had already participated in one or more international large-scale assessment.

Throughout the 1990s, low- and middle-income countries continued to participate in international large-scale assessments: 9 in IEA's Third International Mathematics and Science Study (TIMSS) of 1995 (Beaton, Martin, et al., 1996; Beaton, Mullis, et al., 1996; Martin et al., 1997; Mullis et al., 1997) and 17 in IEA's renamed Trends in International Mathematics and Science Study (also TIMSS) of 1999 (Martin et al., 2000; Mullis et al., 2000). In Latin America, 16 countries participated in the *first Estudio Regional Comparativo y Explicativo* (Regional Comparative and Explanatory Study, or PERCE) of 1998 organised by UNESCO's *Laboratorio Latinamericano de Evaluacion de la Calidad de la Education* (Latin American Laboratory for the Evaluation of Educational Quality, or LLECE) (Ferrer, 2006). In Southern Africa, seven countries joined the Southern African Consortium for Measuring Educational Quality's (SACMEQ) first assessment of 1997 (Makuwa and Maarse, 2013), and in Francophone Africa, the *Programme d'Analyse des Systèmes Educatif de CONFEMEN* (PASEC) exercise was underway in eight countries.[3] By the end of the 20th century, nearly 40 low- or middle-income countries had participated in one or more international or regional large-scale assessment.

Additional countries and assessments, 2000-15

The first decade of the new millennium saw an explosion of interest in international large-scale assessments, including greater participation among low- and middle-income countries:

1. OECD's initial PISA of 2000 (with 7 low- and middle-income countries).
2. PISA 2000 Plus of 2001 (with 8 low- and middle-income countries).
3. IEA's Progress in Reading Literacy Study (PIRLS) of 2001 (16 low- and middle-income countries).
4. SACMEQ II of 2000-2004 (7 low- and middle-income countries).
5. PISA 2003 (14 low- and middle-income countries).
6. TIMSS 2003 (19 low- and middle-income countries).
7. PIRLS 2006 (16 low- and middle-income countries).

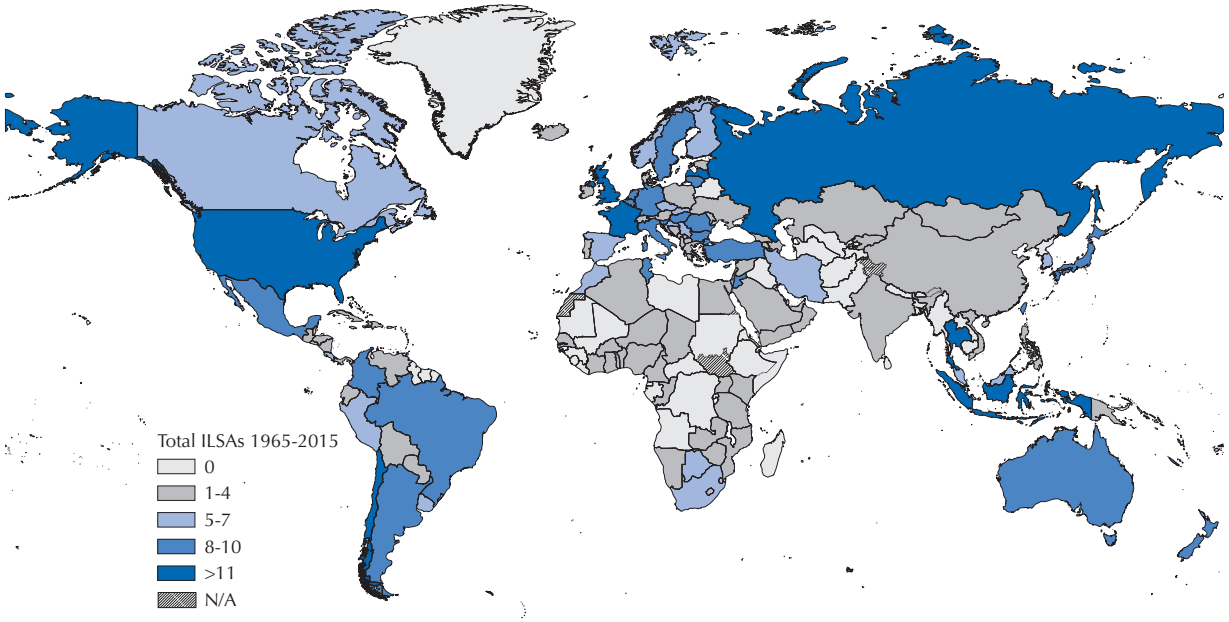
8. PISA 2006 (26 low- and middle-income countries).
9. SACMEQ II of 2006-2011 (14 low- and middle-income countries).
10. Second Regional Comparative and Explanatory Study (SERCE) 2006 (16 low- and middle-income countries).
11. TIMSS 2007 (34 low- and middle-income countries).
12. PISA 2009 and PISA 2009 Plus[4] (34 low- and middle-income countries).

Despite this increase, a very small share of the more than 150 low- and middle-income countries with populations greater than 30 000 participated in these studies. Approximately 22% participated in each of the two major international assessments – PISA 2009/PISA 2009 Plus and TIMSS 2007 – although somewhat different countries participated in the two assessments. Participation rates in regional assessments were higher: 52% of the 31 low- and middle-income Latin American countries participated in SERCE 2006 and 29% of the 48 Sub-Saharan African countries participated in SACMEQ II. In a few cases, countries began participation, often with the support of donors, but either did not complete the assessment or did not report their results internationally.

However, participation in international large-scale assessments has swelled over recent decades. Figure 2.1 shows the number of countries participating in any international large-scale assessment, from 1960 to 2015. The number of participating countries is indicated by ranges from 0 to 13, with the larger number suggesting a sustained assessment culture. The next section of this chapter examines some reasons that low- and middle-income countries participate in international large-scale assessments, with particular reference to PISA.

Figure 2.1

Participation in international large-scale assessments (ILSA), 1965-2015



StatLink http://dx.doi.org/10.1787/888933293835

PISA PARTICIPATION, 2000-15, EMPIRICAL ESTIMATES

Since the first cycle of PISA, the number of countries participating in the assessment has increased by about two-thirds (OECD, 2001, 2004, 2007, 2010, 2014; Walker, 2011). A similar growth in the number of countries participating in other large-scale assessments also occurred over this time period. This increase is largely related to long-standing needs for better education indicators to inform policy. As the United States National Research Council observed in the early 1990s, "The lack of an adequate system of education indicators to inform education policy making has become increasingly apparent" (National Research Council, 1993 as cited in Chabbot and Elliott, 2003). A decade later, Chabbot and Elliott identified PISA specifically as an "indicator" study and observed that a continuing rationale for international studies in education, including large-scale assessments, is to "expand understanding of education as a social and economic institution" (Chabbot and Elliott, 2003).

Lockheed (2013) has argued that another reason for PISA's expansion is the assessment's association with the OECD and the demand generated by economists for valid and reliable cross-national indicators of human capital. As the economist Wößmann notes: "A central pre-requisite for conducting cross-country empirical research is the availability of internationally comparable data. Thus, when investigating the relationship between human capital and the level of economic development across countries, it is crucial to have a reliable measure of human capital which is comparable across countries" (Wößmann, 2002).

Meyer and Benavot (2013) have argued that the rise of PISA is a manifestation of the globalisation of education, broadly, and of educational assessment, specifically. Others have suggested that the growth is due to a largely European regional "peer effect." PISA's growth has also benefitted from the growth in international assessment capacity generally, fuelled in part by previous international and regional assessments and the capacity development activities undertaken as part of both international and national assessments (Kellaghan and Greaney, 2001; Lockheed, 2013).

The OECD acknowledges that the expansion of PISA can be explained by many factors and also accepts that the results of the assessment can be used for several purposes. However, it is important to note that the design, development and implementation of PISA is under the exclusive responsibility of the ministries of education of the participating countries.

Methodology

This chapter tests some of these notions empirically, using a data set covering 214 countries/economies as of 2014 (Lockheed, 2015a). Time series data on country economic status, 1997-2015 (using the four World Bank categories of low-income, lower-middle-income, upper-middle-income and high-income), are matched with information regarding region (East Asia and Pacific, Europe and Central Asia, South Asia, Sub-Saharan Africa, Latin America and the Caribbean, North America, and the Middle East and North Africa), OECD membership, PISA participation, prior participation in IEA's TIMSS,[5] participation in regional assessments (LLECE's PERCE, SERCE and Third Regional Comparative and Explanatory Study [TERCE] in Latin America, three rounds of SACMEQ in Southern Africa, and the most recent PASEC in 2014), and conducting national assessments.[6] Regrettably, complete time-series data on contextual factors, such as the political economy of all these countries, is not available and therefore not included in this analysis.

All countries[7] are included; the assumption is that all countries could have participated in up to six cycles of PISA. The unit of analysis, therefore, is the "country-PISA cycle" or, in aggregate, "participations." Some countries are represented by subnational units only (states, provinces, special administrative regions). This analysis describes some characteristics of the countries that participate in PISA, providing some clues about the determinants of PISA participation.

Increased PISA participation in upper-middle-income and high-income countries

Participation in PISA – defined as completing the assessment and contributing to the international report – has increased from 44 countries/economies in 2000 to 63 countries/economies in 2012, with 74 countries/economies currently participating in the 2015 assessment.[8] Country-level participation appears to have plateaued at around 70 countries/economies. At the same time, the absolute number of participating low- and middle-income countries/economies has increased but has fluctuated considerably over time: 15 in 2000-01, 14 in 2003, 26 in 2006, 34 in 2009-10, 23 in 2012, and 26 in 2015. The OECD reports that several new middle-income countries have confirmed their intention to participate in PISA 2018.

What matters for participation in PISA?

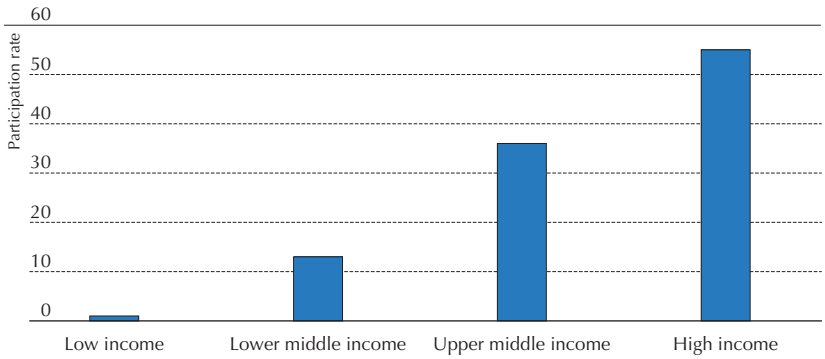
Participation is higher for wealthier countries, countries in Europe and North America, and countries with an established assessment culture, as indicated by evidence of their prior participation in an international large-scale assessment or national assessment.

National wealth matters

The costs associated with PISA participation are more affordable to higher-income countries: 86% of all PISA participations (that is, the total participations across six PISA cycles) come from high-income and upper-middle-income countries, with the overall PISA participation rate[9] of OECD countries (96%) significantly higher than that of non-OECD countries (14%). Moreover, the overall participation rate increases dramatically by income group: the participation rates of high-income countries (55%) and upper-middle-income countries (36%) are substantially higher than those of lower-middle-income countries (13%) and of low-income countries (1.5%). Since 2000, only four countries have participated in PISA at the time they were classified as low-income economies (Albania, India [two states only], Indonesia and Kyrgyzstan).

Figure 2.2

PISA participation rate, by country economic status, 2000-15



Source: Lockheed (2015b), "PISA for Development? The Experiences of Middle-income Countries Participating in PISA", paper presented at the World Bank, June 4, 2015.

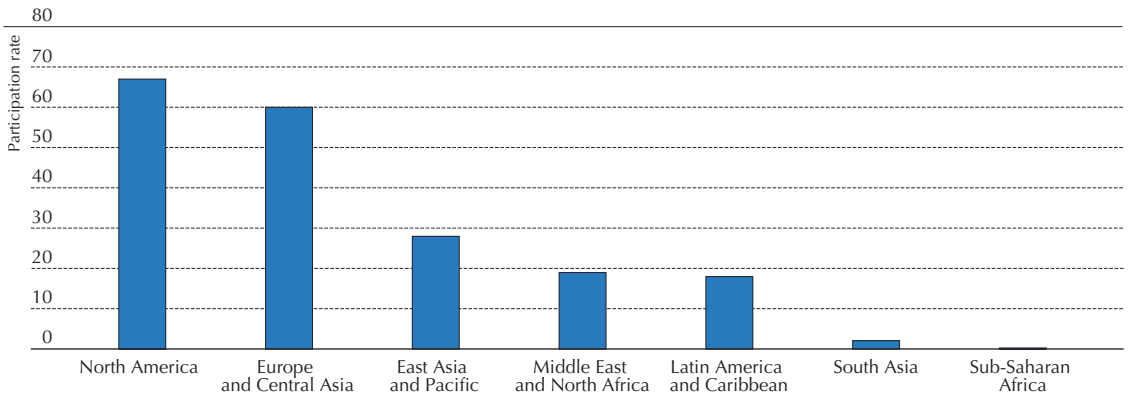
StatLink http://dx.doi.org/10.1787/888933293842

Region matters[10]

PISA participation has spread outwards from OECD countries to their neighbours (Figure 2.3). The overall PISA participation rates of North American countries (67%) and countries in Europe and Central Asia (60%) are higher than participation rates of countries in other regions, with Sub-Saharan Africa recording the lowest participation rate (less than 0.5%). The inclusion of Central Asian countries in the European and Central Asia group hides the fact that the overall participation rate for European Union (EU) countries is almost 100%, whereas the participation rate of non-EU countries in the region is substantially lower.[11]

Figure 2.3

PISA participation rate, by region, 2000-15



Source: Lockheed (2015b), "PISA for Development? The Experiences of Middle-income Countries Participating in PISA", paper presented at the World Bank, June 4, 2015.

StatLink http://dx.doi.org/10.1787/888933293851

Experience with assessments matters

Previous experience with national or international large-scale assessments provides a solid institutional and technical foundation for participation in PISA. The PISA participation rate of countries that have conducted their own national assessments in the three years prior to participating in a PISA cycle is 43%, which is significantly higher than the 15% participation rate of those that have not recently conducted a national assessment. Similarly, the overall participation rate of countries that had participated in TIMSS within five years prior to participating in PISA (64%) is substantially

higher than that of countries that had not participated during this time (15%). This is not surprising; as Eugene Owen, the first chair of the PISA Governing Board, noted: "The stalwarts in IEA were the first countries to participate in PISA" (E. Owen, personal communication, 18 November 2014). Of the 43 OECD and non-OECD countries that participated in PISA 2000 and PISA 2000 Plus, 26 had participated in two or more IEA studies over the previous decade. Participation in a regional assessment, however, appears unrelated to PISA participation.

Odds of participation

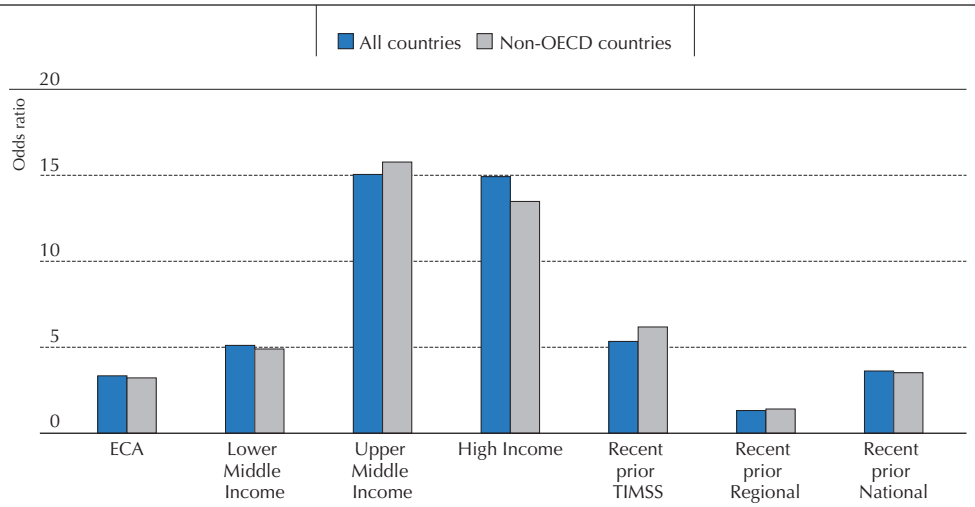
Some of these country characteristics are correlated, such as OECD membership and location in the Europe and Central Asia region. A logistic regression approach provides the opportunity to estimate the odds of participation in any of the six PISA cycles between 2000-15, when all predictors are considered simultaneously.

Overall, the baseline odds of participation in PISA over 15 years are very low: 0.17 participations for every non-participation for all countries, and dropping to 0.02 participations for every non-participation for the low-income group.

The odds ratios reported in Figure 2.4 show the increased odds of PISA participation for each of several country characteristics, taken jointly. These characteristics are: being an OECD country, being a high- or middle-income country at the beginning of each PISA cycle, having previously participated in IEA's TIMSS, having participated in a regional assessment, and having conducted a national assessment in the years prior to a PISA cycle. In all cases, the odds are relative to countries with the following characteristics: non-OECD, non-ECA, low-income and without prior assessment experience. The increased odds from having participated in a regional assessment are not significantly different from zero.

Figure 2.4

Relative odds of PISA participation, by country characteristic, 2000-15



Source: Lockheed (2015b), "PISA for Development? The Experiences of Middle-income Countries Participating in PISA", paper presented at the World Bank, June 4, 2015.

StatLink http://dx.doi.org/10.1787/888933293866

Figure 2.4 summarises results for all countries and for non-OECD countries and shows the following:

1. OECD countries are 105 times more likely to participate than non-OECD, non-Europe and Central Asia low-income countries with no prior assessment experience (not shown).
2. Countries in the Europe and Central Asia region are more than three times as likely to participate as those not in the region, when other things are equal for both groups of countries. In some respects, PISA may be serving as a regional assessment for this group of countries.
3. The likelihood of participation in a PISA cycle increases according to a country's economic status as measured during the first year of the cycle. High- and upper-middle-income countries are about 15 times more likely to participate than low-income countries, other things being equal. This underscores how participation costs act as a deterrent to participation for low- and lower-middle-income countries.

4. Countries that have participated in a recent TIMSS assessment are about five times more likely to subsequently participate in PISA compared with countries that have not recently participated in TIMSS or other assessments. The effects are somewhat greater in non-OECD countries. Prior participation in an international assessment is likely to have provided groundwork for PISA participation.
5. Countries that have conducted a national assessment within three years prior to a PISA cycle are over three times more likely to participate in PISA compared with other countries. National assessments can provide the technical and administrative foundation for participation in an international assessment.
6. Countries that have participated in a regional assessment are no more likely to participate in PISA than other countries. Regional assessments may be substitutes for international assessments, reducing the demand for the broader international comparisons while indicating an interest in learning from others that are more similar regionally and culturally.

Each of these factors are important correlates of participation in PISA and taken together explain a high share of the total variance in participation rates. Only one factor – participation in regional assessments – is not statistically significant and suggests that independent of the other variables, regional assessments do not lead to PISA participation, although they may serve as alternative indicators of system performance. This pattern of correlates is not substantially different for non-OECD countries.

VARIATIONS IN PARTICIPATION: ABSENCE AND RETURN

In general, a stable group of countries participates in PISA, neither skipping cycles nor leaving the assessment. As of 2015, ten countries have returned to PISA after skipping one or two cycles (listed in Table 2.1), and three countries that have not participated since PISA 2009/PISA 2009 Plus have confirmed their intention to participate in PISA 2018: Azerbaijan, Kyrgyzstan and Panama.[12] The countries that have returned after skipping one or two cycles include four high-income countries (Chile, Israel, Malta, and Trinidad and Tobago), five upper-middle-income countries (Albania, Argentina, Bulgaria, Former Yugoslav Republic of Macedonia [hereafter 'FYROM'] and Peru), and one lower-middle-income country (Georgia). Half of these returning countries are from the Europe and Central Asia region, underscoring that region's influence in PISA participation. The OECD considers that there is nothing wrong with a country making selective use of the PISA instruments and is flexible in responding to the cases where a country choses to skip a cycle or returns to the assessment after a gap.

Neither a country's performance level nor change in PISA assessment performance appear to be systematically related to its decision to participate in a specific PISA cycle. All four countries that participated in PISA 2009/ PISA 2009 Plus, but did not participate in PISA 2012 and PISA 2015, performed below the OECD average in the year before dropping out, but whether or not their scores had improved or declined relative to a previous assessment does not seem to have mattered. For example, Kyrgyzstan's scores increased between PISA 2006 and PISA 2009, whereas Azerbaijan's scores declined over the same period. At the same time, many other countries with scores below the OECD average continue to participate.

Table 2.1

Countries returning to PISA after skipping cycles

Country	Skipped PISA Cycle	PISA score (PISA cycle year) prior to skipping			PISA Cycle of return
		Mathematics	Reading	Science	
Argentina	2003	388 (2001)	418	396	2006
Bulgaria	2003	430 (2001)	430	448	2006
Chile	2003	384 (2001)	410	415	2006
Israel	2003	433 (2001)	452	434	2006
Albania	2003, 2006	381 (2001)	349	376	2009
Peru	2003, 2006	292 (2001)	327	333	2009
FYROM	2003, 2006, 2009, 2012	381 (2000)	373	401	2015
Trinidad and Tobago	2012	414 (2009)	416	410	2015
Georgia	2012	379 (2010)	374	373	2015
Malta	2012	463 (2010)	442	461	2015

Note: 2001 and 2010 were PISA Plus assessments.
Source: Author's analysis of OECD, 2001; OECD, 2004; OECD, 2007; OECD, 2010; OECD, 2014; Walker, 2011.
StatLink http://dx.doi.org/10.1787/888933293972

Identifying why a country decides to skip a specific cycle of PISA is difficult as it may be related to any number of factors, including changes in governmental strategic directions, challenges created by the timing of a specific cycle of the assessment, or even the main domain of the assessment. The two-year gap between the PISA 2000 Plus assessment in 2001 and PISA 2003, and between PISA 2009 Plus in 2010 and PISA 2012 may have presented financial and technical challenges that influenced a country's decision regarding whether or not to participate in 2003 or 2012. Of the eight low- and middle-income countries that participated in PISA 2000 Plus, only two (Indonesia and Thailand, both with long histories of participation in international large-scale assessments) participated in PISA 2003, but four more returned for PISA 2006. Similarly, of the seven low-income and middle-income countries[13] that participated in PISA 2009 Plus, only three (Costa Rica, Malaysia and Viet Nam) participated in PISA 2012, but five are participating in PISA 2015.[14]

In some cases, the focus domain for the assessments, which changes on a three-year cycle, may have been important. For example, Albania and Peru participated in PISA 2000 Plus, which focused on reading literacy. These two countries were then absent in 2003 and 2006 but returned for 2009, which also focused on reading literacy.

Governments change and their strategic directions change with them, which can result in intermittent participation, as in the case of Georgia and Trinidad and Tobago, which both skipped PISA 2012 but returned for PISA 2015.

Returning to PISA appears unrelated to education system performance. While all returnees reported scores that were lower than the OECD average, their scores were not dissimilar from those of many other participating countries.

CASE STUDIES OF MIDDLE-INCOME COUNTRIES' INITIAL PARTICIPATION IN PISA

Case studies can provide further details on why countries join PISA. Information for each case study has been compiled from the following: public documents; interviews and/or emails with World Bank and Millennium Challenge Corporation staff, and with National Project Managers; non-restricted World Bank project documents; and OECD studies. Confidentiality with respect to the interviews has been maintained, unless otherwise authorised by the interviewee.

Eight of the nine middle-income countries selected for case studies were chosen on the basis of: *i)* participation in PISA; *ii)* prior participation or non-participation in TIMSS; and *iii)* economic status. The ninth country (Kyrgyzstan) was chosen on the basis of having not participated in PISA since 2009.[15] Georgia, Indonesia, Kyrgyzstan and Viet Nam are all lower-middle-income countries, while Brazil, Bulgaria, Colombia, Jordan and Turkey are upper-middle-income countries. In addition, Bulgaria is a European Union state and Turkey is a member of the OECD. Some basic education indicators for these countries are provided in Table 2.2.

Table 2.2

Basic economic and education indicators, case study countries

Most recent year available

	Lower middle income				Upper middle income				
	Georgia	Indonesia	Viet Nam	Kyrgyzstan	Bulgaria	Brazil	Colombia	Jordan	Turkey
GDP (USD billions)	16.14	868.35	171.39	7.23	54.48	2 245.67	378.42	33.68	822.14
GNI-PPP (USD billions)	31.52	2 315.07	454.98	17.64	110.51	2 955.98	577.82	75.35	1 409.02
Population (millions)	4.48	246.9	89.7	5.7	7.3	198.6	48.3	6.46	74.9
Expenditure on education as % of total government expenditure	6.7	18.1	20.9	18.6	11.2	14.6	15.8	--	--
Public expenditure on education as % of GDP	1.99	3.6	6.3	6.8	4.1	5.8	4.4	--	--
Primary net enrolment rate	98.6	92.2	98.1	90.5	94.9	--	83.9	97.1	94
Secondary net enrolment rate	--	76.1	--	80.4	85.3	--	73.6	87.9	82.1

Source: World Bank EdStats (updated April 2015).

StatLink http://dx.doi.org/10.1787/888933293987

Two-thirds of the case-study countries had experience with international large-scale assessments before they joined PISA. Both Georgia and Indonesia (among the lower-middle-income case study countries) and all countries other than Brazil (among the upper-middle-income countries) had participated in two or more international large-scale assessments prior to joining PISA. The participation of the case study countries, 1990-2015, is summarised in Table 2.3.

The experiences of the case study countries are presented below in two groups: lower-middle-income and upper-middle-income countries.

Table 2.3

Case study countries' participation in international large-scale assessments, 1990-2015

International large-scale assessment	Lower middle income				Upper middle income				
	Georgia	Indonesia	Viet Nam	Kyrgyzstan	Bulgaria	Brazil	Colombia	Jordan	Turkey
IAEP 1991	-	-	-	-	-	yes	-	yes	-
IRLS 1991	-	yes	-	-	-	-	-	-	-
TIMSS 1995	-	yes	-	-	yes	-	yes	-	-
TIMSS 1999	-	yes	-	-	yes	-	-	yes	yes
PISA 2000	-	yes	-	-	yes	yes	-	-	-
PIRLS 2001	-	-	-	-	yes	-	yes	-	yes
TIMSS 2003	-	yes	-	-	yes	-	-	yes	-
PISA 2003	-	yes	-	-		yes	-	-	yes
PISA 2006	-	yes	-	yes	yes	yes	yes	yes	yes
PIRLS 2006	yes	yes	-	-	yes	-	-	-	-
TIMSS 2007	yes	yes	-	-	yes	-	yes	yes	yes
PISA 2009	yes	yes	-	yes	yes	yes	yes	yes	yes
PIRLS 2011	yes	yes	-	-	-	-	yes	-	yes
TIMSS 2011	yes	yes	-	-	-	-	-	yes	-
PISA 2012	-	yes	yes	-	yes	yes	yes	yes	yes
TIMSS 2015	-	-	-	-	-	-	-	yes	yes
PISA 2015	**yes**	**yes**	**yes**		**yes**	**yes**	**yes**	**yes**	**yes**
Total ILSAs	6	13	2	2	11	7	8	10	10

Sources: Author's analysis of Beaton, Martin et al., 1996; Beaton, Mullis et al., 1996; Elley, 1992; Lapointe et al., 1992; Martin et al., 1997; Martin et al., 2000; Mullis et al., 2000; Mullis et al., 1997; OECD, 2014; OECD, 2010; OECD, 2007; OECD, 2004; OECD, 2001; Walker, 2011.

StatLink http://dx.doi.org/10.1787/888933293995

Lower-middle-income countries

Georgia

Georgia's participation in PISA has been encouraged and financed by donors. Georgia is a lower-middle-income country in the European and Central Asia region with an estimated population of around 5 million and a GDP of USD 16.14 billion as of 2013. It boasts near universal literacy, despite its below-average expenditure of 2% of GDP on education. It has received high marks from the World Bank for its anti-corruption measures. Both popular and political support for integration with the West has been high, albeit in flux (Cecire, 2015; World Bank, 2012).

With grant support from the World Bank (through a Development Grant Facility grant to the IEA), Georgia participated in PIRLS 2006 and TIMSS 2007; it also participated in PISA 2009, TIMSS 2011 and PIRLS 2011. Although Georgia prepared for PISA 2012, the assessment was cancelled. Georgia has re-joined PISA for 2015. Georgia had one experience of implementing a large-scale national assessment in 2004 prior to participating in these international large-scale assessments.

With project support from the World Bank, Georgia established the National Assessment and Examinations Center (NAEC) in 2002. Its founding director was a strong supporter of both fair tests and international assessments and served as national project manager for several of the previously mentioned international studies. The main responsibility of NAEC was to conduct university entrance examinations in a fair and equitable way, involving new, tightly controlled implementation procedures with a high degree of security. This reform of the university entrance examination was widely heralded as a success, and the NAEC Director was commended. In addition, NAEC assumed responsibility for conducting international large-scale assessments, and successfully completed the PIRLS, TIMSS and PISA studies between 2006 and 2011.

Although the results of PISA 2009 Plus were not broadly disseminated by the Georgian media, as shown in Chapter 5, they did appear in online blogs and forums. Moreover, results from PISA 2009 Plus were used by one donor – the Millennium Challenge Corporation – during the development of its first education grant to underscore the need for quality and equity improvements in education and to document a widening gender gap in mathematics performance favouring male students in secondary school.

Indonesia

Indonesia's participation in PISA followed its prior participation in earlier IEA studies, its own UNESCO-supported national assessment study in 1978, and its early investment in building technical capacity for educational measurement. Indonesia is a lower-middle-income country of Pacific Asia. With a population of 250 million and a GDP of USD 868.3 billion (as of 2013), its total spending on education is relatively low at 3.6% of GDP. Indonesia has participated in international assessments continuously since 1995,[16] with one of the highest rates of participation overall. A grant from the World Bank, through the Development Grant Facility grant to the IEA, supported Indonesia's participation fees for PIRLS 2006 and TIMSS 2007, and a World Bank Dutch trust fund covered international participation fees for PISA 2012.

Despite not being an OECD member country, Indonesia decided to join PISA to complement its ongoing participation in the IEA studies. One interviewee noted that it joined in order to "learn from both approaches and styles." That is, by joining both the IEA studies and PISA, Indonesia "expected to have hands-on experiences and expertise in the different ways of conducting large scale assessment and to develop [its] own model of a national assessment system."

Indonesia's participation in PISA can also be linked to the country's investment in measurement capacity through overseas doctoral studies in psychometrics in the 1980s and 1990s. In the early 1990s, the Ministry of Education and Culture had an Office of Educational and Cultural Research and Development with a research staff of 236 and an annual operational and development budget of USD 7 million. Within this office was the Centre for Examination System Research and Development, headed by a PhD graduate in statistics from the University of California, Los Angeles (UCLA) (Umar, 1987), who served as the national research co-ordinator for three TIMSS cycles: TIMSS 1995, TIMSS 1999 and TIMSS 2003. Two other PhD graduates in psychometrics (Mohandas, 1996; Hayal, 1992), with expertise in test equating and item banking as well as with experience in analysing TIMSS data, later joined this department and became the national project managers for PISA 2000. Thus, by the time that Indonesia joined PISA, it had considerable experience with the operational complexities of participating in international large-scale assessments, as well as significant technical capacity in psychometrics, which may have made PISA participation more feasible. Today, the National Testing Centre of the Ministry of Education and Culture includes a unit with 22 staff working on international assessments, national assessments and examinations.

Viet Nam

Viet Nam's participation in PISA 2012 benefitted from the country's prior experience with seven national large-scale assessments carried out since 2001, and associated capacity building support received from development partners. Strong central government support that viewed building human capital as essential for economic development and support from donors also helped Viet Nam to participate in PISA. In addition, the Ministry of Education wished to learn from PISA to further develop its own national assessment system. Viet Nam is a lower-middle-income country in East Asia whose economic growth has declined in recent years: its 5% growth in GDP during 2012 was the lowest rate of growth since 1998 (World Bank, 2013). With a population of 90 million and a GDP of USD 171.4 billion, its total spending on education is relatively high at 6.3% of GDP (in 2010). Viet Nam has a 98% primary enrolment rate, a lower secondary enrolment rate above 80%, and an upper secondary enrolment rate close to 60% (Bodewig, Badiani-Magnusson, and Macdonald, 2014). The OECD estimates that only about two-thirds of Viet Nam's 15-year-olds were reflected in PISA 2012 (OECD, 2014).

Following a pilot assessment, 1997-1999,Viet Nam's first national assessment was completed in 2001 by the Viet Nam National Institute for Education Sciences (VNIES), since renamed the Center of Education Quality Evaluation (CEQE). The OECD, UNESCO/International Institute for Educational Planning (IIEP) and the Australian Council for Educational Research helped provide expertise[17] for this assessment (Griffin, 2007; Postlethwaite and Leung, 2007). Although the results from the national assessment were not widely disseminated outside the Ministry of Education, the Minister embraced the results (one interviewee noted that his approach was "tell me the problems and let's figure out how to solve them").

This first national assessment demonstrated significant problems that would need to be solved. Actions were then taken and a second national assessment was conducted in 2007 to help evaluate the effects of the solutions that had been put into place. This report was widely disseminated. Since then, CEQE has carried out five national assessments at various grades and has completed one PASEC assessment (UNESCO, 2015; CONFEMEN, 2015). CEQE is responsible for both national and international assessments and has a specialised unit for PISA that comprises 15 staff. The World Bank has provided support for national assessments, and the Asian Development Bank has provided financial support for both national assessments and PISA participation (Bodewig et al., 2014).

Kyrgyzstan

Donor encouragement and financing has enabled Kyrgyzstan's participation in two PISA cycles. Changes of government and political transitions have contributed to the country's non-participation in subsequent cycles. Kyrgyzstan is a lower-middle-income country in Central Asia, with a population of 5.7 million and a GDP of USD 7.2 billion. GDP growth is high, at 10.5% in 2013, but the country relies on donors to provide support for education reform and innovation. The World Bank's 2009 Systems Approach for Better Education Results (SABER) – Student Assessment Country Report for Kyrgyzstan notes that the country has been conducting national large-scale assessments on an irregular basis, and that those conducted were driven by donor agencies. The country's Monitoring Learning Achievements (MLA) assessment of fourth grade students in mathematics, grammar, and life skills was conducted in 2001 and 2005, with the United Nations Children's Fund (UNICEF) providing funding for the 2005 exercise. The World Bank supported two national large-scale assessments of fourth and eighth grade students in mathematics, natural sciences, and reading comprehension in 2007 and 2009. There are no government policy documents regulating or authorising such assessments. While some ministerial orders have been issued, these are not policy documents.

Kyrgyzstan participated in PISA 2006 and PISA 2009, strongly encouraged by the World Bank and facilitated by grant financing through the Bank's rural education project of 2005. A World Bank report notes that the PISA results from these two cycles were "presented in a consolidated national report and distributed to key decision makers and schools," but that no further steps had been taken to participate in any other international assessments (World Bank, 2009).

The World Bank's Russia Education Aid for Development (READ) Trust Fund programme has played an active role in assessment in Kyrgyzstan, working with the three major assessment organisations: the Centre for Education Assessment and Teaching Methods, the Kyrgyz Academy of Education, and the National Testing Centre. Support for assessment through READ has focused on classroom and national assessments and on improving the technical quality of the national school-leaving examinations. In 2013, READ consultants recommended that Kyrgyzstan not participate in international large-scale assessments at that time, including PISA, on the grounds that PISA was pitched too high for lower-middle-income economies. A recent World Bank project does not include support for any international large-scale assessments, but provides continuing support for enhancing the national assessment (World Bank, 2014).

Upper-middle-income countries

Identifying the reasons for PISA participation among upper-middle-income countries is difficult, since – like Indonesia – many have decades of experience with both national and international large-scale assessments, and the decisions to participate were made two decades ago. In some cases, donor dialogue and support have contributed to decisions regarding participation.

Bulgaria

Bulgaria has participated in the majority of international large-scale assessments since 1995, including the first PISA in 2000. The country has also carried out periodic large-scale national assessments since 1994. Bulgaria is an upper-middle-income country with a population of 7.3 million and a GDP of USD 54.5 billion in 2013. It joined the EU in 2007 and one interviewee noted that: "there is no doubt in anyone's mind that Bulgaria should participate in PISA." In 2005, Bulgaria was a lower-middle-income country, and Bulgaria's participation in IEA's PIRLS 2006 and TIMSS 2007 was supported by a grant from the World Bank, through the Development Grant Facility to IEA to cover the international participation fees. The institution responsible for international large-scale assessments in Bulgaria is the Center for Control and Assessment of the Quality in School Education, which is a unit of the Ministry of Education and Science. Within the Center, a small department of eight professionals deals with PISA, as well as with the country's participation in TIMSS, PIRLS, IEA's International Computer Competence Study (ICCS), and the OECD's Teaching and Learning International Study (TALIS).

Brazil

Brazil first participated in an international large-scale assessment in 1990 (two major cities, Sao Paulo and Fortaleza, participated in the International Assessment of Educational Progress [IAEP]) and Brazil has participated in PISA consistently since 2000. Brazil is an upper-middle-income country, with a population of 198.6 million and a GDP of USD 2 245 billion in 2013. Brazil's National Institute of Educational Studies and Research (INEP) was founded in 1937 and has a staff of 440. It is responsible for: *i)* the national basic education assessment system (SAEB), launched in 1990 for grades 4 and 8, and comprising numerous assessments and school leaving examinations; *ii)* the higher education assessment system that evaluates the quality of undergraduate courses and higher education institutions and revalidates foreign medical diplomas; and *iii)* international programmes, including those sponsored by the *Laboratorio Latinamericano de Evaluacion de la Calidad de la Education* (Latin American Laboratory for the Evaluation of Educational Quality, or LLECE)

and the OECD. The Deputy Minister of Education, formerly the president of INEP, is deputy chair of the PISA Governing Board. Within INEP a small team is responsible for each assessment. In addition, non-government agencies, such as the Carlos Chagas Foundation, provide testing and measurement services for university entrance examinations.

Colombia

Colombia's participation in PISA coincided with the increase in the number of countries participating in international large-scale assessments during the mid-2000s. Colombia is an upper-middle-income country, with a population of 48.3 million and a GDP of USD 378 billion in 2013. Colombia has carried out national large-scale assessments since 1991, and has participated regularly in international large-scale assessments, beginning with TIMSS 1995, the first LLECE study (PERCE) in 1997, and PIRLS in 2001. Since 2006 it has participated in three cycles of PISA, two additional cycles of LLECE assessments (SERCE and TERCE), two cycles of TIMSS and one additional cycle of PIRLS. Participation in TIMSS 2007 was facilitated by a grant from the World Bank through the IEA, but otherwise donors have not supported international assessment activities in Colombia. PISA and the other international large-scale assessments are the responsibility of the Colombian Institute for Educational Evaluation (ICFES), an independent agency that provides assessment and evaluation services to the Ministry of Education, including managing a university selection examination. ICFES has an up-to-date website with information regarding its studies, results and research papers. The website does not list a separate unit for international large-scale assessments.

Jordan

Jordan's participation in international large-scale assessments has been encouraged by international donors and supported at the highest level of government since the late 1980s. Jordan is an upper-middle-income country in the Middle East, with a population of 6.5 million and a GDP of USD 33.7 million in 2013. Jordan was the only country in the region to participate in the IAEP in 1991, supported through grants to Educational Testing Service from the US National Science Foundation, the National Center for Education Statistics, and the Carnegie Corporation (Lapointe, Askew, et al., 1992; Lapointe, Mead, et al., 1992). In the late 1990s, the World Bank and other donors helped finance the establishment of the National Center for Human Resources Development (NCHRD), which has since held responsibility for conducting both national assessments and international large-scale assessments in Jordan. The NCHRD has a staff of about 30 professionals, and a monitoring and evaluation unit that carries out PISA and other studies. During much of its existence, the NCHRD has received special attention from the kingdom's royal family. Jordan has participated regularly in TIMSS since 1999 and in PISA since 2006. A World Bank loan to Jordan in 2003 included funding for participation in TIMSS 2003, TIMSS 2007 and PISA 2006, although ultimately the United States Agency for International Development (USAID) and the United National Development Programme (UNDP) financed the international participation fees and the costs of international travel for the national research co-ordinators.

Turkey

Turkey, the only OECD country among the case study countries, is an upper-middle-income country straddling Europe and Asia, with a population of 75 million and a GDP of USD 822 billion. It has participated in international large-scale assessments since 1999. Turkey's participation in PISA 2003 can be linked to its interest in joining the European Union at the time, and by the fact that all OECD countries were participating in PISA. Turkey made a significant effort to develop its assessment system and to support units within the Ministry of National Education and an organisation that develops and administers university entrance exams. Two World Bank "Adaptable Program Loans" for education included support for the assessment of learning achievement. The first project in 1998 supported the Department of Educational Research and Development (ERDD) of the Ministry of National Education in completing a national assessment in Turkish, mathematics, science and social science; it also provided support for ERDD to prepare national reports from TIMSS 1999 and PIRLS 2001. The second project in 2005 was initially designed to support Turkey's participation in international assessments, including PISA, and to use PISA as a performance indicator. After the project was restructured during implementation, however, "none of the planned … assessments were conducted" and all references to improved student performance were removed from the project's results framework (World Bank, 2012b). Nonetheless, Turkey has continued to participate in PISA and has seen its scores improve.

CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES

Conclusions

Countries participate in PISA for a range of reasons. Wealthier countries participate more often than less wealthy countries, and all OECD countries participate. Some countries with aspirations for membership in the OECD or the EU may participate for political reasons. Countries in Europe and North America participate, while many countries in South Asia, Latin America and Sub-Saharan Africa do not.

In general, money matters. Few low-income countries have ever participated: across 316 possible PISA participations for all countries classified as low-income at any time from 1997 to 2012, only five low-income PISA participations have ever occurred. As a country's economy improves, however, participation in PISA becomes more likely. Lower-middle-income countries have participated when participation is encouraged and financed by donors and when there is a high level of political commitment. The decision to participate in PISA is made at a senior level for all countries, with an official notification provided to the OECD. In many countries, the decision takes the form of a policy document. The participation of upper-middle-income countries is made easier by their long and steady support of related activities, such as national assessments and examinations, and the fact that they have established technical and administrative capacity for conducting international large-scale assessments.

Prior experience – with other international large-scale assessments or with national large-scale assessments – increases a country's likelihood of participating in PISA. This prior experience may have helped to establish sufficient capacity in a country for undertaking large-scale assessments; it also may also have demonstrated the utility of such endeavours. While the PISA cycle itself may build assessment capacity, countries with prior experience in other international large-scale assessments are likely to have established a culture of assessment. Many countries participating in PISA 2015 have a quarter century of relevant experience for doing so.

Beyond these broad generalisations, the country case studies suggest that different countries participate in PISA for different reasons, depending on their needs and capacities. The process of PISA participation begins with an initial signal of intent, which may take place several years before any given PISA cycle, and a five-year commitment to activities before, during and after the main assessment.

Implications for the expansion of PISA participation

The implications of the participation patterns in PISA described above underscore the relatively limited possibility for the expansion of PISA as it currently exists. Assuming that PISA 2015 is completed successfully by all countries listed as participants, this means that one-third of all countries – covering a large share of the world's population of 15-year-olds – will be participants. Expanding further into non-participating high-income countries would involve countries such as Saudi Arabia (a confirmed participant in PISA 2018) as well as the many relatively small non-participating high-income countries, such as Brunei Darussalam (another confirmed participant in PISA 2018), Equatorial Guinea or San Marino. Expanding to non-participating upper-middle-income countries would involve many countries in Latin America, approximately half of which are regular participants in the regional assessments led by the LLECE, and eleven of which already have participated in PISA. Nevertheless, interest in PISA in Latin America may grow. For example, Panama has confirmed its participation in PISA 2018 and Ecuador, Guatemala and Paraguay are participating in PISA for Development.

For upper-middle-income countries

The opportunity for further expansion of PISA appears to be greatest in those few upper-middle-income countries that have previously participated in IEA's TIMSS and/or have conducted regular national assessments but have not yet participated in PISA, for example: Botswana, Iran, Jamaica and South Africa. All of these countries report secondary-level enrolment rates well over 80%, with lower-secondary enrolment rates over 90%, which is important for assessing 15-year-olds in school. All of these countries also have prior experience with national, regional and/or international assessments.

For lower-middle-income and low-income countries

Eight lower-middle-income countries have previously participated in TIMSS and could be considered for inclusion. Two aspects of their country conditions, however, indicate that participation is unlikely: low secondary-level school enrolment rates in some of the economies, and recent conflict in others.[18] Among these countries, only Armenia and Morocco (a confirmed participant in PISA 2018) appear to offer an opportunity for PISA participation.

Lower-middle-income countries that have carried out regular national assessments, such as the Philippines, Senegal and Zambia, may also be considered. The Philippines is a confirmed participant in PISA 2018, and Senegal and Zambia are participating in PISA for Development, which could also stimulate interest in PISA among other countries in the Africa region.

Low-income Sub-Saharan African and South Asian countries appear to offer little opportunity for PISA expansion at present. Moreover, the majority of Sub-Saharan African countries already participate in regional assessments (SACMEQ and PASEC), which may limit their interest in undertaking an international assessment.

For development partners

External financing is a major facilitator of participation in international assessments for lower-middle-income countries. This financing may come from direct grants or from components of loans to governments for education projects. For example, several recent education projects supported by the World Bank either directly supported the country's participation in an international assessment or used international assessment data to justify the project design and/or monitor the project's outcomes.

Between 2003 and 2012, international assessment activities were included in World Bank projects in eleven countries that have participated in PISA: Azerbaijan, Brazil, Bulgaria, Indonesia, Jordan, Kyrgyzstan, Moldova, Montenegro, Peru, Tunisia and Turkey (Liberman and Clarke, 2012). In addition, the World Bank supported international assessments over the same period in another eleven countries that were not PISA participants: Bangladesh, El Salvador, Gambia, Honduras, Lebanon, Mongolia, Nepal, Nicaragua, Saint Lucia, Senegal and Ukraine. Most of these latter countries participated in TIMSS or a regional assessment. Interviews and follow-up emails with 21 National Project Managers of the middle-income countries participating in PISA 2015 indicated that nearly half of those participating in PISA 2015 were enabled to do so by external financing.[19]

Sustained financing across multiple PISA cycles is likely to be essential as the full benefit from participating in most international assessments, including PISA, comes from the opportunity to review changes in performance levels over time. Since PISA emphasises a specific performance domain – mathematics, reading, science – once every nine years, detailed estimates of performance changes can best be made over a much longer time frame than the three-year cycle would indicate. This requirement for sustained financing for national, regional and international assessments is recognised in the development of the Assessment for Learning initiative of the Global Partnership for Education,[20] which aims, in part, to improve financing for the participation of low- and lower-middle-income countries in assessments (Global Partnership for Education, 2015).

Notes

1. The IEA Pilot Study tested reading comprehension, mathematics, science, geography and non-verbal ability.

2. Income classification based on World Bank historical country classification data at the time of participation.

3. Until 2012, however, the PASEC assessment was not designed as an ILSA.

4. PISA Plus was administered in 2010 and included 10 low- and middle-income countries.

5. TIMSS, rather than PIRLS, was selected for this indicator since the TIMSS implementation period included years prior to 2000, when PISA was first implemented.

6. Data on national assessments come from the 2015 UNESCO Global Monitoring Report and were kindly provided by the Director.

7. The terms "country" and "economy" are used interchangeably and refer both to whole countries and countries represented by sub-national units and economies.

8. These lists count as separate economies Hong Kong (China), Macao (China), Puerto Rico, Scotland, Shanghai (China) and Chinese Taipei.

9. The participation rate is computed as the number of cycles in which a country has participated divided by the total number of cycles in which it could have participated (up to six to date).

10. Regional country classifications according to World Bank groups for Europe and Central Asia, East Asia and the Pacific, Middle East and North Africa, Latin America and the Caribbean, South Asia and Sub-Saharan Africa.

11. The EU countries are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain, Sweden and the United Kingdom. After joining the EU, all countries have participated in PISA with the following exceptions: Cyprus in 2006 and 2009, Lithuania in 2009, and Malta in 2006 and 2012.

Notes regarding Cyprus

Note by Turkey: The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue".

Note by all the European Union Member States of the OECD and the European Union: The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

12. Serbia participated in 2012, did not participate in 2015, but is expected to participate in 2018. In addition, two states of India participated in PISA 2009 Plus, but neither India as a whole nor these two states have returned to PISA. Similarly, a sub-national region in Venezuela participated in PISA 2009 Plus, but neither Venezuela nor the region have returned. Mauritius participated in PISA 2009 Plus but has not expressed interest in participating in PISA 2018.

13. Including India (Himachal Pradesh and Tamil Nadu states) and Venezuela (Miranda province).

14. PISA 2009 Plus also included two states in India and one province in Venezuela, none of which participated in PISA 2012 or are participating in PISA 2015.

15. The OECD reports that Kyrgyzstan "plans to participate in PISA 2018".

16. Indonesia administered TIMSS 1995, but was unable to do the steps necessary for its data to be included in the international report for this assessment.

17. Neville Postlethwaite of Hamburg University, Ken Ross of UNESCO/IIEP and Patrick Griffin of Melbourne University.

18. These countries are: Fragile or Conflict Affected: Syria, Ukraine, West Bank and Gaza Strip, and Yemen. Less than 80% secondary enrolments: Ghana, Honduras. Possible PISA participants: Armenia, Morocco (which may participate in 2018).

19. Interviews held October 28-31, 2014 during National Project Managers meetings in Dublin, Ireland.

20. The Global Partnership for Education supports 60 developing countries to ensure that every child receives a quality basic education, prioritising the poorest, most vulnerable and those living in fragile and conflict-affected countries.

Note regarding Israel

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

References

Beaton, A., M. Martin, I. Mullis, E. Gonzales, T. Smith and **D. Kelly** (1996), *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study,* Center for the Study of Testing, Evaluation and Education Policy, Boston College, Chestnut Hill, MA.

Beaton, A., I. Mullis, M. Martin, E. Gonzales, D. Kelly and **T. Smith** (1996), *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study,* Center for the Study of Testing, Evaluation and Education Policy, Boston College, Chestnut Hill, MA.

Bodewig, C., R. Badiani-Magnusson and **K. Macdonald** (2014), *Skilling up Vietnam: Preparing the Workforce for a Modern Market Economy*, World Bank Publications, Washington, DC.

Cecire, M. (2015), "The Kremlin Pulls on Georgia ", *Foreign Policy* (March 3), Slate Division of Washington Post Group.

Chabbot, C. and **E. Elliott** (eds.) (2003), *Understanding Others, Educating Ourselves: Getting More from International Comparative Studies in Education*, National Academies Press, Washington DC.

CONFEMEN (2015), Programme d'Analyse des systèmes Educatifs de la CONFEMEN (PASEC), www.confemen.org/le-pasec/.

Elley, W.B. (1992), *How in the World Do Students Read?*, International Association for the Evaluation of Educational Achievement, Amsterdam.

Ferrer, G. (2006), *Educational Assessment Systems in Latin America: Current Practice and Future Challenges*, Partnership for Educational Revitalization in the Americas (PREAL), Washington DC.

Global Partnership for Education (2015), *Strengthening the Global Partnership's Operational Platform (Workstream 1)*, Board of Directors Meeting paper presented May 22-23, 2015, Seoul, Korea.

Griffin, P. (2007), "Mathematics achievement of Vietnamese grade 5 pupils", *Asia Pacific Education Review*, Vol. 8, No. 2, pp. 233-249, Springer.

Hayal, B. (1992), *A Mathematics Item Bank for Indonesia*, unpublished doctoral dissertation, University of Chicago Department of Education.

Husen, T. (1967), *International Study of Achievement in Mathematics: A Comparison of Twelve Countries,* Vol. 1 and 2, John Wiley and Sons, New York.

Keeves, J. (2011), "IEA: From the beginning in 1958 to 1990 " in C. Papanastasiouu, T. Plomp, and Papanastasiou (eds.), *IEA 1958-2008: 50 Years of Experiences and Memories (Vol. 1)*, Cultural Center of the Kykkos Monastery, Nicosia, Cyprus.

Kellaghan, T. and **V. Greaney** (2001), *Using Asesssment to Improve the Quality of Education*, International Institute for Educational Planning, UNESCO, Paris.

Lapointe, A., J. Askew and **N. Mead** (1992), *Learning Science: Report of the International Assessment of Educational Progress,* Educational Testing Service, Princeton, NJ.

Lapointe, A., N. Mead and **J. Askew** (1992), *Learning Mathematics: Report of the International Assessment of Educational Progress*, Educational Testing Service, Princeton, NJ.

Lapointe, A., N. Mead and **G. Phillips** (1989), *A World of Differences: An International Assessment of Science and Mathematics*, Educational Testing Service, Princeton, NJ.

Liberman, J. and **M. Clarke** (2012), *A Review of World Bank Support for Student Assessment Activities in Client Countries, 1998-2009*, World Bank, Washington DC.

Lockheed, M. (2015a), "Why do countries participate in international large-scale assessments? The case of PISA", *Policy Research Working Papers,* WPS 7447, World Bank, Washington DC.

Lockheed, M. (2015b), *PISA for Development? The Experiences of Middle-income Countries Participating in PISA*, paper presented at the World Bank, June 4, 2015.

Lockheed, M. (2013), "Causes and consequences of international assessments in developing countries" in H.-D. Meyer and A. Benavot (eds.), *PISA, Power and Policy: The Emergence of Global Educational Governance*, pp. 163-185, Symposium Books, Oxford.

Makuwa, D.K., and **J. Maarse** (2013). "The Impact of Large-Scale International Assessments: A Case Study of how the Ministry of Education in Namibia Used SACMEQ Assessments to Improve Learning Outcomes", *Research in Comparative and International Education*, Vol. 8, No. 3, pp. 349-358, Sage.

Martin, M., I. Mullis, A. Beaton, E. Gonzales, T. Smith and **D. Kelly** (1997), *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study*, Center for the Study of Testing, Evaluation and Education Policy, Boston College, Boston.

Martin, M., I. Mullis, E. Gonzalez, K. Gregory, T. Smith, S. Chrostowski, R. Garden and **K. O'Connor** (2000), *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*, International Study Center, Boston College, Chestnut Hill, MA.

Meyer, H.-D. and **A. Benavot** (eds.) (2013), *PISA, Power and Policy: The Emergence of Global Educational Governance*, Symposium Books, Ltd, Oxford, UK.

Mohandas, R. (1996), *Test Equating, Problems and Solutions: Equating English Test Forms for the Indonesian Junior Secondary School Final Examination Administered in 1994*, unpublished doctoral dissertation, Flinders University of South Australia.

Mullis, I., M. Martin, E. Gonzalez, K. Gregory, R. Garden, K. O'Connor, S. Chrostowski and **T. Smith** (2000), *TIMSS 1999 International Mathematics Report Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade,* International Study Center, Boston College, Chestnut Hill, MA.

Mullis, I., M. Martin, A. Beaton, E. Gonzales, D. Kelly and **T. Smith** (1997), *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study,* Center for the Study of Testing, Evaluation and Education Policy, Boston College, Boston.

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

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

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

OECD (2001), *Knowledge and Skills for Life: First Results from PISA 2000*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264195905-en.

Postlethwaite, T.N. and **F. Leung** (2007), "Comparing educational achievements", *Comparative Education Research*, pp. 215-239, Springer.

Postlethwaite, T.N. and **D. Wiley** (1992), *The IEA Study of Science II: Science Achievement in Twenty-three Countries*, Vol. 2, Pergamon Press, Oxford.

Robitaille, D.F. and **R.A. Garden** (eds.) (1989), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics*, Pergamon Press, Oxford.

Umar, J. (1987), *Robustness of the Simple Linking Procedures in Item Banking using the Rasch Model*, unpublished doctoral dissertation, UCLA Division of Social Research Methodology.

UNESCO (2015), *Education for All 2000-2015: Achievements and Challenges*, UNESCO, Paris.

Walker, M. (2011), *PISA 2009 Plus Results: Performance of 15-year-olds in Reading, Mathematics and Science for 100 Additional Participants,* ACER Press, Victoria.

Wößmann, L. (2002), "Cross-country evidence on human capital and the level of economic development: the role of measurement issues in education", *Historical Social Research,* Vol. 27, No. 4, pp. 47-76, GESIS.

World Bank (2015), World Bank EdStats.

World Bank (2014), *Kyrgyz Global Partnership for Education*, The World Bank, Washington, DC.

World Bank (2013), *Taking Stock : An Update on Vietnam's Recent Economic Developments*, The World Bank, Washington DC.

World Bank (2012), *Fighting Corruption in Public Services: Chronicling Georgia's Reforms*, The World Bank, Washington, DC.

World Bank (2009), *SABER-Student Assessment Country Report: Kyrgyz Republic*, Washington, DC.

PISA FOR DEVELOPMENT

3

What have been the challenges facing middle-income countries participating in PISA?

This chapter explores three sets of challenges facing middle-income countries that participate in the Programme for International Student Assessment (PISA): costs of participation, capacity for implementing PISA, and the country context for PISA. It reviews the evidence for each type of challenge, drawing on the literature and interviews with representatives from middle-income countries participating in PISA 2015. The first section of the chapter explores the general challenges of cost, as well as the challenges of specific expenditure items, and the role of donor support in meeting the financial challenges. The second section examines a range of technical challenges, including analytical challenges such as instrument development, scientific sampling and data analysis, as well as operational challenges presented by common timelines, sampling in federal countries, translation in multiple languages, survey implementation, and coding. The third section describes context challenges due to a country's political economy, regulatory environment and national culture.

INTRODUCTION

Prior research on the challenges facing countries involved in international assessments suggests three broad areas of deterrents: costs, capacity and context (Bloem, 2015; Lockheed, 2010). Most of the evidence regarding these deterrents is informal, arising from self-reports and casual interviews. Costs involve the direct incremental costs of participation, including international participation fees, travel costs for required meetings, printing and other costs associated with data collection and processing, as well as the personnel costs associated with conducting the assessment. Capacity involves the institutional and technical capacity to conduct the assessment and to analyse the results; this involves national survey implementation, translation demands in countries with a diversity of languages of instruction, and analytical capacity to gain insight from the results (Bloem, 2015). Context involves the broad characteristics of nations that can affect their capacity to implement PISA and their responses to and use of the PISA results. This chapter reviews these challenges, illustrates them with the experience of middle-income countries participating in PISA 2015, and makes suggestions for actions to address these challenges in the future.

FINANCIAL CHALLENGES

Although many countries mentioned that meeting the costs of participation in PISA was a challenge, many of the National Project Managers could not provide an estimate of the actual cost of PISA participation. They were, however, aware of the cost items that presented difficulties in financing and reported how donors had helped to offset these costs.

Costs of PISA

Direct costs of participation in an international large scale assessment are very small when expressed as a share of total expenditure on education (Wolff, 2007). These costs include the minimum international participation fee (set at EUR 182 000, payable at EUR 45 500 per year for four years, for PISA 2015), costs of travel and subsistence to mandatory meeting and training events (which require the use of a country's foreign exchange account), and the in-country costs associated with implementing the assessment. Wolff estimated the costs of participation – including the international participation fee – for PISA 2000 for Peru and PISA 2003 for Uruguay (Wolff, 2007). For Peru, the cost was equivalent to 0.04% of the country's annual expenditure on secondary education, and for Uruguay the cost was equivalent to 0.08%. A World Bank consultant estimated that the cost of PISA 2006 participation for Kyrgyzstan was USD 300 000, including international participation fees (M. Mertaugh, personal communication January 7, 2015). Wagner, Babson and Murphy (2011) report the costs of PISA 2009 as 0.002% of Mexico's annual expenditure on secondary education; for Panama it was 0.04%.

Interviews with representatives from the one low-income country and from 15[1] of the 25[2] middle-income countries participating in PISA 2015 revealed a wide range of estimates for the direct costs of participating in PISA, as well as an absence of information about costs in some cases. Many of the direct costs of PISA are related to the overall sample size, which was around 5 000 to 6 000 students in the middle-income countries participating in PISA 2012, with little variation among countries except for Mexico (33 806 students), Brazil (20 091) and Colombia (11 173). While translations into multiple languages could also increase costs, this was not mentioned as a major cost driver.

Information about the cost of PISA 2015 was available for only eight middle-income economies, and Costa Rica and Kosovo reported the costs of the field trials only (Table 3.1). Peru mentioned the high costs associated with computer-based assessment in PISA 2015. In all cases for which education expenditure data were made available, the reported costs of PISA amounted to less than one-half of one percentage point – and in some cases much less – of total expenditure on secondary education.

Participation in PISA involves financial commitment to a set of prescribed requirements, as laid out in the agreement documents and the National Project Manager's manual. To provide participating countries with more complete information regarding costs, the OECD could consider including additional information about the level of effort required, both human and financial, in the National Project Manager's manual.

For PISA 2015, approximately 90 separate tasks are required to complete the seven-step process of participation. In addition, the PISA National Project Manager's manual includes detailed information regarding communicating with the four main contractors responsible for leading all aspects of PISA 2015 implementation, including communication standards, which are broken down into seven "core" areas: cognitive assessment framework development; development of platform for the computer-based assessment; instrument development, scaling and analysis; survey operations and procedures; sampling; background questionnaire framework and development; and oversight and management.

Table 3.1

Direct costs for PISA 2015 implementation, selected economies

Country	Cost of field trial	Cost of main data collection	Cost of reporting	Total cost	As % of annual expenditure on secondary education
Indonesia*	USD 80 000	USD 200 000	USD 30 000	USD 310 000	0.029
Peru	--	USD 500 000	--	USD 2 000 000	0.105
Thailand	--	--	--	USD 979 600	0.01
Brazil*	--	--	--	USD 2 000 000	0.004
Jordan	--	--	--	USD 141 000	0.026
Bulgaria*	--	USD 22 040	--	--	0.003
Costa Rica	USD 98 000	--	--	--	--
Kosovo	USD 90 000	--	--	--	--

* Case study country.

Sources: Author's interviews with National Project Managers for PISA 2015, October 28-31, 2014 in Dublin, Ireland; UNESCO Institute for Statistics (2015) for annual expenditures on secondary education computed from government expenditure per student and total secondary student enrollment.

StatLink http://dx.doi.org/10.1787/888933294008

Although the National Project Manager's manual explains the tasks and communication requirements in considerable detail, it only provides explicit guidance regarding the required skills and level of effort needed and the expected incremental costs for a few of these tasks (e.g. the level of effort needed for coding open-ended responses and occupational categories is provided.) Providing an estimate of the level of effort needed, along with a description of the skills and competencies of persons required for various tasks, could be valuable for countries that have no or little experience with international large-scale assessments and may not accurately assess the resource needs related to each task. PISA for Development has been working with participating countries to develop a capacity needs analysis, a capacity building plan[3] and a project implementation plan that include detailed budgets for implementation in each country (for more details see Box 4.1 in Chapter 4).

Challenges of specific expenditure items

One-third of the countries interviewed reported that they had no difficulty in obtaining funding for PISA 2015 or paying for any expenditure item. All the remaining countries, including all the case study countries other than Colombia, identified at least one challenge in financing PISA. The lower-middle-income case study countries reported more challenges than were reported by upper-middle-income countries (Table 3.2).

Table 3.2

Financial challenges noted by case study countries, PISA 2015

	Lower middle income			Upper middle income			
Financial challenges	Georgia	Indonesia	Viet Nam	Bulgaria	Brazil	Colombia	Jordan
Salaries (full- and part-time)	-	Yes	Yes	-	Yes	-	-
Office space	-	-	Yes	-	-	-	-
Office supplies	-	Yes	Yes	-	Yes	-	-
Travel (international and local)	-	Yes	Yes	-	-	-	Yes
PISA fees	Yes	Yes	Yes	Yes	-	-	Yes
Reporting/dissemination	-	Yes	Yes	-	-	-	-
Research	Yes	Yes	Yes	Yes	Yes	-	Yes
Printing	-	-	Yes	-	-	-	-
Data processing	-	-	Yes	Yes	-	-	-

Note: Turkey did not send a representative to the National Project Managers Meeting in Dublin, November 2014, and did not respond to emails requesting information, and is therefore omitted from this table.

Source: Author's interviews with National Project Managers for PISA 2015, October 28-31, 2014 in Dublin, Ireland.

StatLink http://dx.doi.org/10.1787/888933294019

Most countries interviewed reported that they had experienced difficulties in finding funding for selected expenditure items. The most frequently mentioned items were: research and development related to PISA, international participation fees, salaries or honoraria for part-time local staff or experts working on PISA, and international travel. Part-time salaries and honoraria are important for survey administration, coding and data entry as these are short-time tasks that do not

require full-time staff. In Europe and North America such staff are often teachers whose summer holidays coincide with the PISA schedule, but this is not the case in all countries. In some countries, "topping up" the salaries of civil servants – including teachers – is not permitted, although in other countries salary "premiums" are given for specialised work, which could apply to these assessment tasks.

International travel required for National Project Manager meetings and training was also singled out as particularly challenging, in one case it required prior approval from the country's president. Other expenditure items that were mentioned included the costs associated with printing materials for training, local transport, and salaries for full-time staff working on PISA. Countries that had implemented the computer-based assessment identified specific financial challenges, particularly the absence of simple processes for contracting outside services, such as renting computers.

Donor support for PISA

Some lower-middle-income countries have received commitments from international donors to support their participation in PISA cycles (Table 3.3 reports commitments from the World Bank to support international assessments, including PISA). For example, the World Bank financed the participation of Kyrgyzstan in PISA 2006 and 2009. Donors also helped finance the PISA 2015 participation of some middle-income countries. Interviewees from four of the five lower-middle-income economies (Georgia, Kosovo, the Republic of Moldova and Viet Nam) said that their country would not have been able to participate in PISA 2015 without donor support for the assessment. The only lower-middle-income country reporting that costs of PISA 2015 were entirely covered by the country's internal funding sources was Indonesia. In addition, two upper-middle-income countries (Jordan and Lebanon) said that donor support for the international participation fees, and often for PISA-related travel costs, had enabled their participation. All other upper-middle-income countries interviewed reported that the costs of PISA were covered by the country's internal funding sources.

Table 3.3

Recent World Bank projects supporting PISA and other international large-scale assessments

Project	Year project approved	International large-scale assessment mentioned for support
Senegal – Quality and Equity of Basic Education	2013	PIRLS, PISA, TIMSS
Peru – Basic Education Project	2012	PISA, TERCE
Moldova – Education Reform Project	2012	PISA
Sri Lanka – Transforming the school education system as the foundation of a knowledge hub project	2011	PISA, TIMSS
Nicaragua – Second support to the education sector project	2011	TERCE
Lebanon – Second Education Development Project	2010	PISA, TIMSS
Jordan – Second Education Reform for the Knowledge Economy Project	2009	PISA, TIMSS

Note: Senegal is participating in PISA for Development. Sri Lanka has not participated in PISA.
Source: World Bank (2015a), Project database, http://go.worldbank.org/KTPE1WKU20.
StatLink http://dx.doi.org/10.1787/888933294026

TECHNICAL CHALLENGES

The technical challenges for participating in an international large-scale assessment are numerous and can be divided into two broad categories: analytical challenges and operational challenges. Analytical challenges include challenges of psychometrics, sampling (for sample-based assessments), data analysis and reporting. Operational challenges include data collection (including challenges related to data security and access to data) and processing. These challenges occur during each of the seven distinct stages of the assessment process: a first review phase (including instrument development), the field trial preparation, field trialling of instruments and processes, reviewing the field trial, the main study preparation, the main study implementation, a final review stage, and reporting.[4]

The steps for implementing a PISA cycle have been developed and refined over nearly two decades of experience, and the detail provided may be considered "best practice". For countries and contractors, the steps involved in completing a PISA cycle are similar to the steps involved in completing any standardised assessment, including national assessments (Greaney and Kellaghan, 2008), but there are a few differences, as described in the following sections. The technical challenges identified in case study countries are summarised in Table 3.4.

Table 3.4

Technical challenges noted by case study countries, PISA 2015

	Lower middle income			Upper middle income			
Technical challenges	Georgia	Indonesia	Viet Nam	Bulgaria	Brazil	Colombia	Jordan
Item writing	-	-	Yes	-	-	-	-
Translation	Yes	-	Yes	-	-	-	-
Sampling	Yes	Yes	-	-	Yes	-	-
Access to schools	-	-	-	-	Yes	-	-
Platform for CBA	-	-	-	Yes	-	-	-
Administering PISA	Yes	Yes	-	-	Yes	-	-
Training coders	Yes	Yes	-	-	-	-	-
Coding	Yes	-	Yes	-	-	-	-
Submitting data	-	-		-	Yes	-	-
National report	-	Yes	Yes	-	-	-	-

Note: CBA is computer-based assessment.

Source: Author's interviews with National Project Managers for PISA 2015, October 28-31, 2014 in Dublin, Ireland.

StatLink http://dx.doi.org/10.1787/888933294035

Analytical challenges

What distinguishes the analytical dimensions of PISA and other international large-scale assessments from national assessments is the degree to which contractors have responsibility for three of the most highly technical aspects of assessment: the development and refinement of the assessment instruments, scientific sampling and multi-country data analysis. While countries are involved at varying levels, the centralisation of these technical procedures in the hands of international contractors ensures that they are carried out comparably across countries and conform to high quality professional standards.

Framework and instrument development

PISA, along with other international large-scale assessments, invites participating countries to review conceptual frameworks and submit and review new items, including training for item writing and coding. Expert groups and item developers from contractors, such as the Australian Council for Education Research or the Educational Testing Service, provide the bulk of the materials, which are then revised, commented on and validated by participating countries through a series of consultations and meetings. The contractors retain responsibility for the psychometric reliability and validity of items, including assessing item bias, maintaining an item bank and ensuring that assessment instruments can be equated to maintain score comparability across countries and over time. This process is described fully in a recent paper that identifies issues in adapting the cognitive instruments for use in a wide range of country contexts (Adams and Cresswell, 2014).

Scientific sampling

Other contractors, such as Westat,[5] are responsible for the scientific sampling of the schools and students who represent the country in the assessment. Sampling for PISA is somewhat more complex than for other international assessments, since the PISA sample is age-based (involving students in multiple grades), whereas others are grade-based (and can sample intact classrooms). Countries are responsible for providing up to date lists of schools and students, sample stratification (for example, by instructional language or sub-national unit), and for using the sampling software provided, but these tasks are operational rather than analytical.

Data analysis

The majority of data analysis is carried out by the contractors and the OECD, which prepares the international reports. Preparation and dissemination of a national report is the country's responsibility, and this task appears to be a challenge for the middle-income countries participating in PISA. The OECD does not routinely collect country reports and therefore does not include reports from low- or middle-income countries on its website. However, interviews suggest that national reports were prepared and disseminated in about half of the low- and middle-income countries that had participated in PISA 2012. Several countries said that a lack of adequate resources made it difficult to produce a national report. Few countries reported having prepared policy briefs or reports targeted at schools and teachers, although many countries prepared brochures describing the country's results.

Operational challenges

Four operational dimensions of PISA were mentioned as specific challenges by National Project Managers. The first two are dimensions of PISA that distinguish it from typical national standardised assessments: *i)* the need to agree to and maintain a time line with deliverables that are co-ordinated with those of other countries; and *ii)* the additional steps required

for translation of source documents in the official OECD languages (English and French) into one or more languages of instruction and verification of the translated instruments within the agreed timelines. The second two operational challenges of PISA are common to other large-scale assessments, including national assessments: *i)* survey implementation, including gaining access to lists of schools and students for scientific sampling; and *ii)* coding of constructed-response items.

Timelines and schedules, translations, survey implementation and constructed-response coding are required in regional assessments (such as the Regional Comparative and Explanatory Study [ERCE] and the Southern African Consortium for Measuring Educational Quality [SACMEQ]) and other major international assessments (principally, the Trends in International Mathematics and Science Study [TIMSS] and the Progress in International Reading Literacy Study [PIRLS]). These activities may present challenges to some countries joining PISA.

The specific technical challenges faced by low- and middle-income countries participating in PISA are discussed in the following paragraphs. Countries with prior PISA experience and with strong national assessment or examination systems reported fewer operational challenges than countries lacking this experience.

Common timelines

Several countries mentioned the importance of adhering to the timelines for implementation as outlined in the National Project Manager's manual and accompanying documents. One interviewee mentioned that "technical procedures provided by PISA and OECD should be strictly followed", and another emphasised that countries should "do things on time; contact the contractors and do what Core 1-6 says!" Maintaining timelines may be challenging when the PISA timeline conflicts with national or other international assessments or examinations, and other operational tasks may be challenging when appropriate staffing is unavailable. Contractors also mentioned that while most countries participating in PISA 2015 had been able to complete all the tasks required for the field test, the contractors needed to remind some National Project Managers on a near-weekly basis of the actions that needed to be taken to meet pending deadlines.

Sampling

Sampling is a challenge in large federal countries (such as Brazil and Mexico) and in countries with high degrees of decentralisation (such as Indonesia). It is also a challenge when central agencies do not have access to comprehensive lists of schools, and may need to work through sub-national authorities. Many countries lack the education management information systems that would facilitate the sampling task. Decentralised responsibilities for education can also present a challenge with respect to gaining access to schools in a timely manner.

Translation

PISA provides "source documents" in the two official languages of the OECD: English and French. Countries that do not use either as the language of instruction are responsible for translating the survey instruments from these source documents into the language, languages or regional versions of some languages (such as Spanish, Portuguese, French and Arabic) in which the students will be tested. This process requires an understanding not only of the core assessment domains but also specific education terminology in order to maintain clarity and comparability. The process of translation and back translation, often by multiple translators (given the multiple instruments that need to be translated), was mentioned as a challenge in several countries, since this process often required more time than originally scheduled.

To help meet this challenge, the Latin American countries participating in PISA 2015 hired a translation firm to create "source documents" in Spanish. Support for translating the science items was provided by a donor, *Groupo de Analisis para el Desarrollo* (GRADE) in Peru (S. Cueto, personal communication 4 April 2015). Translation challenges may increase if PISA expands further. Many countries that do not currently participate in PISA use multiple languages of instruction, for example, different states in India have different instructional languages. Multiple languages add to the costs and operational complexities, as the instruments must correspond to the source documents and to each other. Translation is also an issue for countries that conduct national assessments in multiple languages, but the "source documents" are more likely to be in a national rather than international language.

Survey implementation

The many operational challenges of PISA and other international large-scale assessments can be offset by detailed instructional manuals. For PISA 2015, the National Project Manager's manual is 177 pages of specific instructions and checklists of activities that cover the four-year development and implementation cycle for the assessment. It is designed to ensure that the results are comparable across countries. Throughout the PISA cycle, National Project Managers are required to organise and carry out approximately 90 separate and related tasks, including attending multiple meetings and training courses, such as: four National Project Managers meetings, international training for the field trial (covering

survey operations, coding and data entry), and international training for the main study (survey operations, paper item coding, and data entry training). These multiple tasks have been grouped into 21 categories for capacity-building activities to better suit low- and middle-income countries (OECD, 2014a; 2014b; 2015a; 2015b).

Countries using the computer-based assessment platform for the first time identified it as a challenge for survey implementation. Issues with the computer-based platform included difficulties in obtaining (renting or purchasing) sufficient computers for the assessment, and scheduling the computers for assessment sessions.

Coding

Coding of constructed-response items requires a high degree of inter-coder reliability for the results to be meaningful. Obtaining this reliability requires considerable training and quality control. Coder training sessions are included in the PISA implementation process, and these sessions are generally well attended: 20 of the 26 middle-income countries participating in PISA 2015 attended coder training provided in 2014. However, coding of open-ended responses appears to take more time than anticipated in the schedule for PISA, particularly when this task requires hiring and training part-time staff. Some countries do not have smooth procedures for obtaining temporary professional staff, such as coders, and the time frame within the PISA schedule is constrained for this activity.

In addition, cultural expectations regarding coding may hinder the process. Specifically, to assess inter-coder reliability (and therefore have the opportunity to discard unreliable coders), double coding open-ended questions (constructed-response items) without resolution of discrepancies is required. This may present problems in countries where the resolution of discrepancies in real time (as, for example, with "table scoring" which involves two or more coders that sit at a table, score the same tests and resolve discrepancies in real time) is a common practice, particularly in the case of scoring national examinations. Contractors identified countries that needed to improve their coding for the main study from the submission of PISA 2015 field trial data that did not meet the inter-coder reliability goal of 92% agreement. It is anticipated that the inter-coder reliability for the main study will meet this goal as a result of continued training.

The OECD recognises these challenges and the National Project Manager's manual is intended to address them. Countries that have participated in several cycles report fewer challenges than those that are relative newcomers to PISA.

CONTEXT CHALLENGES

The context for undertaking a PISA cycle is shaped by a country's political economy, regulatory environment and culture. Over 90% of PISA implementation has taken place in the developed world. As PISA expands its reach to include more non-OECD and more non-Western countries, differences in the context for assessment will need to be taken into account.

Context can affect various aspects of PISA participation in different ways. The political economy can affect changes in administration and education ministers – and hence support for the assessment – as well as the environment for public discourse regarding education reform in general and PISA results in particular. A country's regulatory environment can affect the ease of doing business, particularly in terms of contracting for short-term goods and services. And cultural differences among countries (or among regions) can shape the nature of the cross-cultural contacts that are essential for both adherence to agreed-upon schedules and peer learning.

These broad generalisations were underscored by several interviewees, who said that building an enabling environment was a key challenge to participating in PISA. Issues identified included the challenge presented by changes in administration and consequently education ministers, difficulties encountered by independent assessment agencies that have weak linkages to the financial arms of the ministry, and procedural challenges in obtaining financing. In addition, some interviewees mentioned that outreach to stakeholders was essential and that it was important to "inform everybody about the assessment", "target other stakeholders, especially families, to raise their awareness about the study and help them understand the results", and to introduce PISA to teachers "to help them realise the importance of PISA." The OECD has noted that countries have shared their methods for outreach and knowledge dissemination, such as brochures, flyers and websites, with other countries adopting these approaches.

Political economy

The political economy of most long-standing PISA participants can be characterised as "democratic", meaning that the countries typically have a high degree of citizen voice and accountability, as measured by perceptions of various stakeholders. The World Bank Governance Indicators project recently ranked over 200 countries/economies according to their rating on "voice and accountability". This ranking showed that only one (Costa Rica) of the middle-income countries participating in PISA 2015 was rated in the top quartile of all countries for "voice and accountability", compared with 70% of the participating high-income economies (World Bank, 2015a).

Since governments join PISA, government effectiveness can be important for effective implementation. The "governance effectiveness" index from the World Bank's Governance Indicators project reflects the perceptions of the quality of public services; the quality of the civil service and the degree of its independence from political pressures; the quality of policy formation and implementation; and the credibility of the government's commitment to such policies (Kaufmann et al., 2010; World Bank, 2015b). Among the middle-income countries participating in PISA 2015, only one (Malaysia) is ranked among the top quarter of the more than 200 countries with respect to government effectiveness. By comparison, 72% of the participating high-income countries are ranked in the top quarter. The effectiveness of the public institutions responsible for PISA may also influence the implementation of the assessment.

Freedom of the press affects how easily the results from PISA can be shared broadly with the public. The World Press Freedom Index provides a measure of press freedom that has been used by the European Union (EU) in discussions of the European Neighbourhood Policy in 2014, and is also a useful index for considering the opportunity for public discussion of PISA results in middle-income countries (European Commission, 2015). The index ranked 180 countries/economies on "the degree of freedom that journalists, news organisations and netizens enjoy in each country, and the efforts made by the authorities to respect and ensure respect for this freedom" (Reporters Without Borders, 2014). Only one country (Costa Rica) of the 25 middle-income countries participating in PISA 2015 is among the top quarter of all 180 countries rated, and 30% of middle-income PISA 2015 participants are in the lowest quarter. By comparison 65% of the 46 high-income countries/economies participating in PISA 2015 are in the top quarter and only two are in the lowest quarter. The presence, or absence, of press freedom may help explain the results regarding the public dialogue of PISA results, discussed in Chapter 5.

Regulatory environment

A country's regulatory environment – a key feature of "ease of doing business" – can affect operational aspects of PISA, such as contracting needed temporary staff (for coding, data entry, etc.), purchasing supplies and equipment, or leasing computers for computer-based assessment. The World Bank has ranked 189 countries for overall ease of doing business, with specific reference to a country's regulatory environment for operating a local firm. This regulatory environment may also affect the local agencies, including ministry units, responsible for conducting PISA. PISA implementation will be easier in countries with a more favourable environment for doing business. A higher share of countries participating in PISA are in the top quartile for ease of doing business, in comparison with all countries. Specifically, 75% of high-income countries participating in PISA 2015 and 35% of middle-income countries are in the top quartile, compared with 59% of all high-income countries, and 12% of all middle-income countries (World Bank, 2014).

National culture

National culture can be difficult to assess, but research over 50 years has shown that differences in national cultures can affect the ease of cross-cultural communication in business contexts (Hofstede et al., 2010). Since communication among peers and between countries and contractors is essential for implementing PISA (and other international large-scale assessments), awareness of differences in national culture may be needed, and adjustments made, to avoid miscommunication.

One aspect of international large-scale assessments that can be misunderstood is their status as "low-stakes" assessments. In some countries, the students and schools selected for participation may believe, incorrectly, that the results will affect them personally, thus turning a "low-stakes" assessment into a "high-stakes" assessment in their eyes. This in turn can increase test preparation activities in an effort to improve scores.

Furthermore, sensitivities regarding national performance in PISA can influence the degree to which the results are discussed publicly. Such sensitivities may create a political risk for elected governments, which need to defend the quality of their education systems. The response to the results from international large-scale assessments in different countries indicates how cultural differences can affect behaviour. In 2000, Germany suffered "PISA shock" when the country scored below the OECD average and in the middle of the score distribution. This "PISA shock" was a call to action that has been well documented (Dixon et al., 2013; Ertl, 2006). By comparison, when Mexico scored below expectations in TIMSS 1995, it declined to publicly release its scores (J. Prawda, personal communication, 5 January 2015).

CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES

Conclusions

Middle-income countries that are currently participating in PISA face financial, technical and contextual challenges in conducting the assessment. Efforts are already under way to bring additional countries into PISA and these will need to take these challenges into account.

1. The main financial challenge is adequate and sustained financing – over a 10- to 15-year period – for the units or agencies responsible for PISA, so that countries can participate in the multiple rounds of the assessment that are needed to measure trends in performance.
2. The main technical challenges are largely operational ones: translation into multiple languages from source documents provided in only a few languages, sampling (particularly in large federal or decentralised systems), and selected implementation aspects as described above. Although only a few middle-income countries have developed national reports or have conducted secondary analysis of the PISA data, this was identified as a financial issue rather than lack of technical capacity.
3. The main contextual challenges of the political, regulatory and cultural environment may be overly broad to address through PISA. However, differences between countries that are participating in PISA and those that are not – in terms of government effectiveness, press freedom, ease of doing business and subtle cultural characteristics – may affect PISA implementation.

Implications

For middle-income countries

To address the substantial financial challenges identified by the middle-income countries participating in PISA, countries could:

1. Ensure that budgets for the national units responsible for international assessments, including PISA, are sufficient to cover their costs, and work with development partners to help mitigate financial challenges.
2. Expand resources for research using the PISA data to enable country-specific analyses that could help guide quality and equity improvements relevant to the country.

To address technical challenges, countries could:

1. Ensure that technically relevant representatives of national units responsible for international assessments attend the various capacity building sessions offered by PISA partners and contractors.
2. Consider establishing or expanding units responsible for international assessments, in line with the technical level of effort needed to complete the various tasks involved.

To address contextual challenges, countries could:

1. Analyse their own governance, regulatory and cultural context to identify issues that could compromise the successful implementation of PISA and other international assessments.
2. Improve outreach to stakeholders so that the entire education system – from the central ministry to teachers to parents – has a better understanding of the assessment, what it shows and what it can be used for.

For development partners and the OECD

To address financial challenges and building on the PISA for Development initiative, partners and the OECD could:

1. Develop a cost manual aligned with the National Project Manager's manual that would help countries estimate the actual costs of conducting the assessment. Such a manual could elaborate the level of effort required to complete each task in terms of time, technical qualifications and skills of the required staffing, physical resources needed and associated costs.
2. Establish research and development programmes, possibly through fellowship and grants, that could lead to analyses of PISA data that is more attuned to the policy needs of the individual participating countries.

To address technical challenges and building on the PISA for Development initiative, partners and the OECD could:

1. Provide support for creating additional source documents in major international languages, such as Spanish or Chinese, which could help mitigate some of the translation issues.
2. Provide support for extending peer learning on effective ways to gain co-operation from subnational units, particularly in large federal systems, which could help mitigate some of the sampling issues.

To address contextual challenges and building on the PISA for Development initiative, partners and the OECD could:

1. Consider adjusting certain PISA processes, while maintaining quality standards, as such adjustments might be relevant to specific countries.

Notes

1. Interviews with Brazil, Bulgaria, Colombia, Costa Rica, Mexico, Peru, Thailand (computer-based assessments for PISA 2015), Former Yugoslav Republic of Macedonia, Georgia, Indonesia, Jordan, Kazakhstan, Kosovo, Lebanon, FYROM, Moldova, Viet Nam (paper based assessments for PISA 2015).

2. Lower middle-income economies: Former Yugoslav Republic of Macedonia, Georgia, Indonesia, Kosovo, Moldova, Viet Nam. Upper-middle-income countries: Albania, Algeria, Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Dominican Republic, Jordan, Kazakhstan, Lebanon, FYROM, Malaysia, Mexico, Montenegro, Peru, Romania, Thailand, Tunisia, Turkey.

3.These reports are available at the PISA for Development website: www.oecd.org/pisa/aboutpisa/pisafordevelopment.htm. The Project Implementation Plans are working documents for participating countries and have therefore not been published.

4. Westat is an employee-owned research corporation consulting in statistical design, data collection and management, and research analysis work.

References

Adams, R. and **J. Cresswell** (2014), *PISA for Development Technical Strand 2: Enhancement of PISA Cognitive Instruments*, Australian Council for Education Research.

Bloem, S. (2015), "PISA for Low-and Middle-income Countries", *Compare: A Journal of Comparative and International Education,* Vol. 45, No. 3, pp. 481-486. Taylor and Francis.

Dixon, R., C. Arndt, M. Mullers, J. Vakkuri, K. Engblom-Pelkkala and **C. Hood** (2013), "A Lever for Improvement or a Magnet for Blame? Press and Political Responses to International educational rankings in four EU countries", *Public Administration,* Vol. 91, No. 2, pp. 484-505. John Wiley and Sons.

Ertl, H. (2006), "Educational Standards and the Changing Discourse on Education: The Reception and Consequences of the PISA Study in Germany", *Oxford Review of Education,* Vol. 32, No. 5, pp. 619-634, Taylor and Francis.

European Commission (2015), *Implementation of the European Neighbourhood Policy: Statistics*, Staff Working Document, European Commission, Brussels.

Greaney, V. and **T. Kellaghan** (2008), *National Assessments of Educational Achievement, Vol. 1: Assessing NationalAachievement Levels in Education*, World Bank Publications, Washington DC.

Hofstede, G., G.J. Hofstede and **M. Minkov** (2010), *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and its Importance for Survival*, John Wiley and Sons, New York.

Kaufmann, D., A. Kraay and **M. Mastruzzi** (2010), *The Worldwide Governance Indicators: Methodology and Analytical Issues*. Brookings Institution, Washington, DC.

Lockheed, M. (2010), *The Craft of Education Assessment: Does Participating in International and Regional Assessments Build Assessment Capacity in Developing Countries?*, International Association for the Evaluation of Educational Achievement, Amsterdam.

OECD (2015a), *Capacity Needs Analysis Report: Guatemala,* PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Guatemala%20CNA%20reportFINAL_SL.pdf.

OECD (2015b), *Capacity Needs Analysis Report: Ecuador*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Ecuador%20CNA%20ReportFINAL_SL.pdf.

OECD (2014a), *Capacity Needs Analysis Report: Senegal*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Pisa-D_CNA_Senegal_FINAL_revised.pdf.

OECD (2014b), *Capacity Needs Analysis Report: Zambia*, PISA for Development. OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf.

Reporters Without Borders (2014), *World Press Freedom Index 2014*, Reporters Without Borders, Paris, France.

UNESCO Institute for Statistics (2015), www.uis.unesco.org.

Wagner, D., A. Babson and **K. Murphy** (2011), "How Much Is Learning Measurement Worth? Assessment Costs in Low-Income Countries", *Current Issues in Comparative Education,* Vol. 14, No. 1, pp. 3-21. Teachers College, Columbia Univerity, New York.

Wolff, L. (2007), *The Costs of Student Assessments in Latin America*, PREAL, Washington DC.

World Bank (2015a) Project data base, http://go.worldbank.org/KTPE1WKU20.

World Bank (2015b), *Worldwide Governance Indicators*, World Bank, Washington DC.

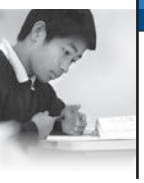
World Bank (2014), *Doing Business 2015: Going Beyond Efficiency*, The World Bank, Washington DC.

PISA FOR DEVELOPMENT

4

What have been the capacity-building outcomes for countries participating in PISA?

In order for countries to carry out more effective assessments of student learning, they must build or enhance country capacity in three areas: the enabling context for assessment, the alignment of all aspects of the education system, and assessment quality. The first section of this chapter examines how capacity has been measured, drawing on World Bank Systems Approach for Better Education Results (SABER)-Student Assessment reports. The second section explores how development partners and multilateral donors have helped build capacity for assessment, through training, technical publications and hands-on capacity building. The third section examines the evidence linking these activities to improved capacity, focusing on evidence from various international large-scale assessments and the Programme for International Student Assessment (PISA). This section also explains the capacity development activities of PISA for Development.



INTRODUCTION

In order for countries to carry out more effective assessments of student learning they must build capacity in three key areas: enabling context, system alignment and assessment quality (Clarke, 2012). The enabling context includes: setting clear policies for assessment, having strong public engagement for assessment, having regular funding for assessment, having strong organisational structures for assessment and having effective human resources for assessment. System alignment includes: aligning the assessment with system learning goals, curricula, textbooks and teacher development. Assessment quality includes: ensuring the technical quality of the assessment (from design to implementation to analysis and reporting) and effective use of the results.

Within each of these broad categories there are many specific capacity elements that involve building institutional, organisational and individual capacities. This chapter explores the extent to which participation in PISA has led to improved capacity in two of these broad areas: the enabling context and the technical quality of assessments. System alignment issues are more specific to a particular country and are best addressed in relation to national and sub-national assessments that are linked to a country's official learning goals and curricula.

MEASURING ASSESSMENT CAPACITY

Systematic empirical measurements of a country's capacity for learning assessment are available for relatively few countries. More often, capacity is described in general terms. For example, one report reviews in detail the educational assessment systems of 22 Latin American countries with respect to 7 categories of analysis and 27 distinct themes, but provides no systematic summary of the results (Ferrer, 2006).

To obtain a systematic measurement of assessment capacity, The World Bank's Systems Approach for Better Education Results (SABER)-Student Assessment programme has assisted more than 20 countries to carry out reviews of their assessment capacity. These countries cover five regions (Europe and Central Asia, the Middle-East and North Africa, sub-Saharan Africa, South Asia and East Asia) and include three of the case study countries (Jordan, Kyrgyzstan, and Viet Nam). Similarly, the OECD and the World Bank have supported detailed capacity needs analyses for the six countries participating in the PISA for Development pilot programme, four of which have been completed (OECD, 2014a, 2014b, 2015a, 2015b). Both the SABER-Student Assessment and capacity needs analysis exercises evaluated countries on a specific set of capacities and then rated them as "latent", "emerging", "established" or "advanced" in each capacity area.

Results from these two systematic approaches show considerable need for capacity development in the area of large-scale assessments, with most of the ratings falling into the "emerging" category. Among the 20 countries with published SABER-Student Assessment evaluations, none received an "advanced" rating for either national or international large-scale assessments. A rating of "established" was given to 10% of countries for national large-scale assessments and to 20% of countries for international large-scale assessments. A rating of "emerging" was given to 70% of countries for national large-scale assessments and to 45% of countries for international large-scale assessments. Finally, a rating of "latent" was given to 20% of countries for national large-scale assessments and to 33% of countries for international large-scale assessments.

— Table 4.1 —

National and international large-scale assessment capacity ratings

	Jordan (draft as of 2014)		Kyrgyzstan (as of 2009)		Viet Nam (as of 2009)	
	NLSA	ILSA	NLSA	ILSA	NLSA	ILSA
Overall rating	Emerging	Emerging	Emerging	Emerging	Emerging	Latent
Enabling context	Emerging	Emerging	Emerging	Emerging	Emerging	Latent
System alignment	Emerging	Established	Emerging	Emerging	Established	Emerging
Assessment quality	Emerging	Emerging	Emerging	Emerging	Established	Latent

Notes: Ratings are based on the numerical average for the categories rated on each dimension, ranging from 1 = latent to 4 = advanced.
National large-scale assessment (NLSA) and international large-scale assessment (ILSA).
Source: SABER-Student Assessment website and country reports saber.worldbank.org/index.cfm.

StatLink http://dx.doi.org/10.1787/888933294048

SABER-Student Assessment ratings for the three areas of enabling context, system alignment and assessment quality are generally similar within a country. For example, two of the case study countries for which SABER-Student Assessment evaluations are available were rated as having "emerging" capacity for both national and international large-scale assessments across most of the three broad areas (Table 4.1). However, differences in capacity for undertaking national and international large-scale assessments were found for all categories in Viet Nam, and for system alignment in Jordan. At the time of the SABER-Student Assessment evaluation, Viet Nam had not yet participated in either the Trends in International Mathematics and Science Study (TIMSS) or PISA.

BUILDING CAPACITY FOR LARGE-SCALE ASSESSMENTS

Donors have provided support to low- and middle-income countries for assessment capacity building in three main ways: *i)* short-term and advanced degree level training; *ii)* technical publications on assessment; and *iii)* "hands-on" assessment capacity building. Most of this support has been targeted at raising awareness and improving the quality of assessments. In addition, support for secondary analysis of data is provided by most of the sponsoring organisations, including the OECD.

Short-term and advanced degree training

Throughout the 1990s, the World Bank Institute offered courses designed to build or enhance a country's capacity for carrying out learning assessments (Murphy et al., 1996). From 2002 to 2007, the World Bank sponsored regional workshops on national assessments, first in Asia between 2002 and 2004 (Bangkok, Colombo, Hanoi and Vientiane), and then in Africa between 2004 and 2006 (Accra, Addis Ababa, Cairo, Kampala, Maseru and Pretoria). In 2007, the Asian and African participants were brought together to share their experiences and receive additional training on item analysis (V. Greaney, personal communication, 11 July 2015).[1]

In 2014, the World Bank and the International Monetary Fund (IMF) sponsored a three-day workshop on national and international assessments for ministry of education officials from 20 countries in the Middle East and North Africa, and the United Nations Education, Science and Culture Organization (UNESCO) organised a four-day workshop on large-scale assessments for the Network on Education Quality Monitoring in Asia-Pacific. These workshops were intended to raise awareness about large-scale learning assessments and their utility for education policy.

Universities, professional associations and assessment institutions provide learning opportunities relevant to PISA participants through advanced degree training and technical short-term courses related to large-scale assessments. Universities provide individuals with the opportunity to deepen their knowledge regarding the technical issues in large-scale assessments through post-graduate work in statistics and psychometrics. Bilateral and multi-lateral donor agencies have often financed the participation of graduate students from low and middle-income countries in these programmes.

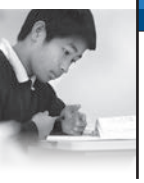
Professional associations, such as the American Educational Research Association, also offer short term courses related to the analysis of international large-scale assessments. These courses are typically offered concurrently with the associations' annual meetings. Scholarships have been provided by these associations to participants from low-income and lower-middle-income countries.

Not-for profit institutions have provided fee-based courses on data analysis. The Australian Council for Education Research (ACER) has offered courses on analysing PISA data. The Educational Testing Service (ETS) has offered courses on analysing the results from the Programme for the International Assessment of Adult Competencies (PIAAC). The International Association for the Evaluation of Educational Achievement (IEA) has provided training in secondary analysis of TIMSS data (Wagemaker, 2014). Since 2007, the IEA-ETS Research Institute (IERI) Academy has offered 20 courses on topics such as assessment design, item-response theory, Bayesian models, and population modelling. It has also provided scholarships for participants from low-income and lower-middle-income countries.

Technical publications on assessment

Publications raise awareness and provide technical information regarding learning assessments. During the 1990s, the World Bank supported publications designed to raise awareness of the importance of learning assessments, particularly national assessments of learning (Kellaghan and Greaney, 2001; Murphy et al., 1996). The World Bank's SABER-Student Assessment programme is intended to raise awareness of various types of student assessments.

More recently, the World Bank has supported the preparation of a series of five books intended to build capacity for the multiple activities involved in undertaking, analysing and using national learning assessments (Box 4.1).



Box 4.1 **Five World Bank books designed to build capacity for national assessments**

- *National Assessments of Educational Achievement, Volume 1: Assessing National Achievement Levels in Education* (Greaney and Kellaghan, 2008).
- *National Assessments of Educational Achievement, Volume 2: Developing Tests and Questionnaires for a National Assessment of Educational Achievement* (Anderson and Morgan, 2008).
- *National Assessments of Educational Achievement, Volume 3: Implementing a National Assessment of Educational Achievement* (Greaney and Kellaghan, 2012).
- *National Assessments of Educational Achievement, Volume 4: Analyzing Data from a National Assessment of Educational Achievement* (Shiel and Cartwright, 2015).
- *National Assessments Of Educational Achievement, Volume 5: Using the Results of a National Assessment of Educational Achievement* (Kellaghan, Greaney, and Murray, 2009).

Publications by other not-for-profit organisations have also covered awareness and capacity building for assessment in middle-income countries. For example, the Partnership for Educational Revitalization in the Americas (PREAL) has published reviews of large-scale assessment programmes in Latin America (Ferrer, 2006). The IEA, in co-operation with ETS, sponsors an open-access journal, *Large-Scale Assessments in Education*, designed to "contribute to the science of large-scale assessments, help disseminate state-of-the-art information about empirical research using these databases and make the results available to policy makers and researchers around the world." (Large-scale Assessments in Education, 2015). The American Psychological Association (APA) lists over 50 journals in the English language that focus on various technical aspects of non-clinical educational assessments (APA Division 5, 2011). PISA's National Project Manager's manual, and the technical reports prepared for each PISA cycle, can also be considered capacity building publications.

Hands-on assessment capacity building

Direct assessment capacity building has been provided to low- and middle-income countries in two ways: through financial support from donors for participation in international large-scale assessments, and through the on-going technical support accompanying participation in large-scale assessments.

Donor support has enabled low- and middle-income countries to participate in regional and international assessments, with participation often including direct training activities (Lockheed, 2013). For example, the Association for the Development of Education in Africa promoted capacity development for assessment through its partnership with the Southern African Consortium for Measuring Educational Quality (SACMEQ). UNESCO, through the International Institute for Educational Planning (IIEP), supported the first two cycles of SACMEQ and, through the Latin-American Laboratory for the Assessment of the Quality of Education (LLECE), the first two cycles of the Latin American assessment: the Regional Comparative and Explanatory Study (ERCE). The United Nations Development Programme (UNDP) assisted five Arab states to participate in TIMSS 2003. The World Bank, through its Development Grant Facility, supported 18 low- and middle-income countries' participation in TIMSS 1999, 24 low- and middle-income countries' participation in the Progress in International Reading Literacy Study (PIRLS) 2001 and/or TIMSS 2003, 20 middle-income countries' participation in PIRLS 2006 and/or TIMSS 2007, and also provided support to SACMEQ and the *Programme d'Analyse des Systèmes Educatif de CONFEMEN* (PASEC).

Participating countries typically receive hands-on training when participating in an international large-scale assessment. This type of training is often operational, but enables participants to build skills relevant to the assessment. For example, at the 2014 National Project Managers' meeting for PISA 2015, held in Dublin, Ireland, specialised "hands-on" training and individual consultations related to use of the sampling software for the assessment were provided.

Donors have also supported countries to prepare national reports, and researchers to conduct secondary analysis of PISA data. For example, the Analysis for Development Group (GRADE) in Peru helped with the preparation of the national report for PISA 2012 in Peru.

Support for secondary analysis of data

Three types of support for secondary analysis have been provided by donors and sponsors of large-scale international assessments: high quality data and data documentation, fellowships for data analysis and grants for conference

participation. Sponsors of international large-scale assessments facilitate secondary analysis through providing high quality data and data documentation for researchers. Technical reports on the assessments are prepared and published on the websites of the IEA, the OECD, the LLECE and SACMEQ, among others. Reports typically include the questionnaires, codebooks and analytic programmes (typically in Statistical Analysis Software – SAS – or Statistical Package for the Social Sciences – SPSS) to facilitate analysis. In addition, clean data sets are also made available. In the case of SACMEQ advance registration is required to use the data.

Fellowships for data analysis have also been provided. For example, in 2012 the OECD launched the Thomas J. Alexander fellowship programme to encourage researchers to use OECD education data to study issues of quality and equity. By 2015, 21 researchers – including fellows from Brazil, Chile, the People's Republic of China (hereafter 'China'), Poland, South Africa and Turkey – had been awarded fellowships to work alongside OECD analysts on issues of education quality and equity.

Grants for conference participation have been used to stimulate research on education quality in low- and middle-income countries, using data from international large-scale assessments. For example, in 2005, PREAL, in collaboration with the World Bank and the Global Development Network (GDN), provided grants for researchers from developing and transition countries to participate in a conference on educational research in these countries. Nearly 200 full papers were submitted, subjected to peer review, and critiqued by expert researchers.[2] The highest quality papers were subsequently published (Lockheed and Cueto, 2006).

RESULTS FROM CAPACITY DEVELOPMENT ACTIVITIES

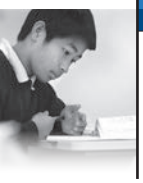
Few international large-scale assessment capacity development activities for low- or middle-income countries have been formally evaluated. Formal evaluations have been conducted for IEA and SACMEQ capacity development activities, but not for PISA. The SABER-Student Assessment evaluations cannot be used to determine whether or not PISA participation has built assessment capacity as they were only conducted once, typically after a country had already participated in an international large-scale assessment, including PISA, so any change in assessment capacity due to PISA participation is not observable.

Moreover, the overlap between participation in PISA and TIMSS makes it difficult to disentangle the effects of PISA from those of TIMSS.[3] However, countries that have participated in international large-scale assessments received higher capacity ratings. For example, the three countries – Kazakhstan, Serbia and the United Arab Emirates – that had participated in both TIMSS and PISA prior to the SABER-Student Assessment evaluation received an "established" rating (3.0) for international large-scale assessment capacity. Also, the average rating for international large-scale assessment capacity was higher (2.5) for the 8 countries that had participated in either or both of the two international large-scale assessments than it was for the 10 countries that had not participated (1.4). These associations between participation and capacity ratings suggest that international large-scale assessment participation may have enhanced these countries' assessment capacity. Alternatively, countries with existing assessment capacity may have been more likely to participate in international large-scale assessments.[4]

Capacity development through participation in TIMSS, PIRLS and SACMEQ

Independent evaluations of IEA and SACMEQ activities, required by donors, provide some empirical evidence of impact (Elley, 2002; Gilmore, 2005; Lockheed, 2010; Murimba, 2005). These evaluations, however, did not address all three dimensions of capacity (enabling context, system alignment, assessment quality) and generally focused on assessment quality. The evaluations often concluded that participating in an international large-scale assessment built technical and operational skills for assessment for the professionals who participate in the assessment. The evaluations are limited from a methodological perspective as they are only based on the participants' self-reports of impact.

Nevertheless, participants report positive outcomes. For example, Elley (2002) reported that "In the course of the TIMSS-R project, training sessions were held for national research co-ordinators (NRCs) and members of their team in such technical matters as sampling, test development, coding, data processing and the like, and most NRCs claimed to benefit a great deal from these sessions." Similarly, Gilmore, in assessing the impact of participation in another IEA study (TIMSS/PIRLS) reported that "NRCs were asked to rate the special or on-the-job training that they had received during the projects. Without exception, all NRCs reported that the training in sampling, test development, questionnaire development, data management, and quality control were 'very helpful' or 'helpful'" (Gilmore, 2005). Another evaluation of an IEA programme found that the IEA training and capacity-building activities were viewed as having contributed to institutional and individual capacities related to assessment operations (Lockheed, 2010). Finally, Murimba (2005) noted that participation in SACMEQ's capacity-building initiatives in sampling, data collection and data analysis were positively evaluated.



Capacity development through participation in PISA

PISA has the potential to build or enhance capacity for all dimensions of capacity development: enabling context, system alignment and assessment quality. Systematic evidence that participation specifically in PISA has improved a country's assessment capacity is limited for three reasons: *i)* no formal evaluation has been carried out; *ii)* opting into PISA can be based on a country's existing capacity; and *iii)* the focus of most PISA training has been on ensuring that the assessment was carried out in accordance with existing standards.

PISA's impact on assessment capacity has not yet been evaluated

No formal evaluation of PISA's impact on assessment capacity has been carried out. While the OECD notes that it has helped countries successfully complete the assessment, it has not conducted or sponsored an empirical evaluation of PISA's impact on countries' assessment capacity, particularly for the middle-income participants. There is some anecdotal evidence that PISA has contributed to improvements in the technical quality of assessments.

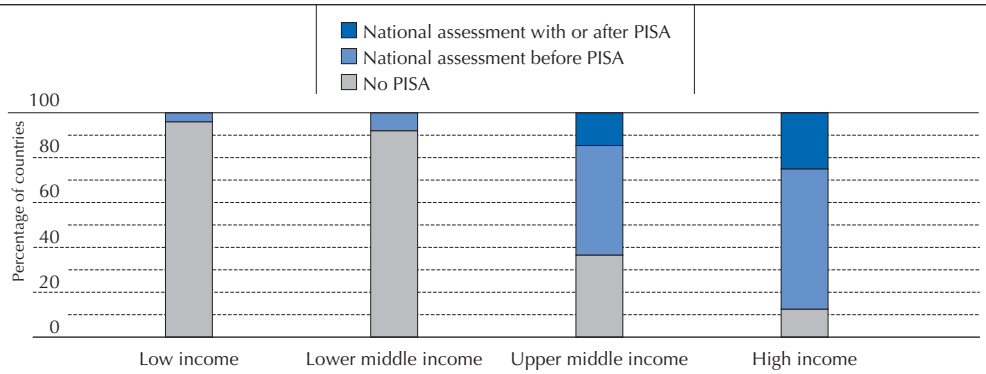
The OECD expects countries to have adequate assessment systems

Initially, the OECD assumed that countries participating in PISA were capable of conducting the assessment according to OECD's standards and therefore did not measure the assessment capacity of the original participants. However, as more non-OECD countries applied to join PISA, the OECD has asked countries to demonstrate that they have the capacity to undertake the assessment, through, for example, the successful completion of a pilot assessment or field test in a PISA cycle. Participation in a prior international large-scale assessment, or evidence of a high quality national assessment, can also demonstrate capacity. The detail of PISA technical standards allows countries to self-select when applying for PISA participation, based on their understanding of these standards. Countries without sufficient existing individual or organisational technical capacity may choose not to apply.

Most of the countries that have participated in PISA have been high-income countries with relatively well-developed examination and assessment systems, and with considerable technical expertise in assessment. Participating countries also often have significant prior experience with other international and regional large-scale assessments, as well as national assessments, as demonstrated in Chapter 2.

Figure 4.1

National assessments build capacity for PISA



Source: Author's analysis of data from Lockheed (2015), "Why do countries participate in international large-scale assessments? The case of PISA", *Policy Research Working Papers*, WPS 7447, World Bank, Washington DC.

StatLink http://dx.doi.org/10.1787/888933293872

Among the 145 countries that UNESCO (2015) identified as having national learning assessments, a high share of first-time PISA participants had several years of experience with national learning assessments. This was particularly the case for upper-middle- and high-income countries (Figure 4.1). About 70% of PISA participants, including 24 low- and middle-income countries, joined PISA after having conducted a national assessment. However, 16 upper-middle-income countries participated in PISA without having had prior national assessment experience.

PISA capacity development focused on technical quality

The capacity development activities of prior PISA cycles have focused on ensuring that the assessment met the standards specified in the technical manuals. As a consequence, training focused on specific technical quality aspects of the assessment and was not designed to enhance the enabling context for assessment or directly contribute to system alignment.

Training has been closely embedded in the semi-annual meetings of the National Project Managers. Changes in the enabling contexts of participating countries may have occurred, however, and may be observed in the sector plans of some middle-income countries. For example, the sector plans of Cambodia refer to goals of performance on and participation in international assessments, and specifically mention PISA and international standards for learning. The evidence regarding PISA's impact on education policy, including agenda setting, is discussed in Chapter 5.

There is some evidence that the training activities related to PISA have built technical quality for assessment in countries. Interviews undertaken for this report with the PISA management teams from 16 low- and middle-income countries participating in PISA 2015 suggest that the PISA experience helped build capacity for conducting better-quality assessments through training related to specific operational tasks. The majority of the country teams, 12 out of 16, recalled that one or more member of the team had received training in selected areas. In response to the question: "At the time your country last participated in PISA, was your PISA unit or team offered opportunities to receive special or on-the-job training related to PISA?" the teams said that someone from the team had received training related to sampling (nine countries), data collection (nine countries), data entry (eight countries), professional scoring of constructed-response items (seven countries), and translation (seven countries).[5] Whether or not these training activities built assessment capacity is not known, but the responses indicate that training raised awareness about many of the technical elements of assessment.

While participation in PISA may have contributed to a country's assessment capacity development, no systematic evidence of this impact is available. Working with contractors and the OECD, countries opt into participating in PISA only after having determined that their assessment capacity is sufficient to undertake the assessment.

Building assessment capacity through PISA for Development

The PISA for Development project includes specific activities designed to enhance assessment capacity, particularly the technical quality of the assessment. The OECD observes that the number of National Project Manager meetings has increased from six to eight, and their duration has increased from four to five days, compared to the main PISA NPM meetings. In addition, the final two meetings will be devoted to analysis topics and developing national reports. As with PISA, the meetings for PISA for Development will be held in participating countries, which facilitates the greater participation of operational staff rather than just the National Project Managers. Larger numbers of participants will also increase the likelihood of reaching a 'critical mass' of experienced staff that renders institutions more resilient to the effects of individual promotion, retirement and other migrations. Whether or not participation in the PISA for Development initiative will change the enabling context for assessment remains to be seen. A steering committee of stakeholders, often formalised as part of a PISA implementation plan, can extend the base of support for the assessment and may improve the enabling context in some countries. Because the capacity needs analysis includes indicators of the enabling context – setting clear policies for assessment, regular funding for the international large-scale assessment unit, and human resources – the impact of PISA participation on the enabling context for assessment can be addressed.

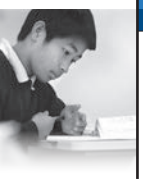
CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES

Conclusions

Participants of international large-scale assessments perceive that participation improves a country's capacity for conducting high quality assessments. However, much remains to be learned about the impact of PISA participation on assessment capacity development in middle-income countries. In most participating countries, PISA implementation has benefitted from an existing assessment capacity that has been developed through participation in other assessments, particularly other international-large scale assessments. The experience of OECD countries suggests that participation in PISA raises awareness of technical standards and improves the quality of country-level PISA reports.

To date, however, there has been no rigorous evaluation of the impact of PISA participation on assessment capacity development in middle-income countries, and the available evidence is limited and mostly anecdotal. Some middle-income countries report that participation has built operational capacity for assessments and that they have introduced new techniques into their ongoing national assessments.

PISA for Development offers an opportunity to obtain empirical evidence regarding the effect of participation in PISA on assessment capacity, since participating countries undertake a capacity needs analysis prior to participation (Box 4.2). This allows countries to identify dimensions of assessment capacity that need strengthening; it also informs the OECD about areas needing enhanced training efforts. A follow-up capacity needs analysis could allow for a comparison of the country's assessment capacity before participating in PISA for Development with its capacity following participation.



Box 4.2 Evaluating PISA's capacity building impact

The PISA for Development initiative offers an opportunity to evaluate the impact of PISA participation on a country's assessment capacity. Baseline indicators of assessment capacity, which are crucial for assessing change, have been established through capacity needs analyses (CNAs) that have been completed in six of the seven pilot countries. The CNAs are structured around three levels of capacity: enabling context, organisation and individual. A software application has been developed that enables a systematic assessment of the country's capacity to carry out PISA for Development (Cartwright (2015). This CNA application produces 117 ratings of specific capacity needs and has been successfully used in Cambodia, Ecuador, Guatemala, Paraguay, Senegal and Zambia. To assess the enabling context capacity needs (48 rating categories), the CNA application uses the SABER-Student Assessment (Clarke, 2012) rating categories for more than 50% of the categories. To assess organisation capacity needs (41 ratings) and individual capacity needs (28 ratings) the CNA application developed rating categories drawn from specific project requirements. Only six SABER-Student Assessment ratings categories are used for organisational capacity and only one is used for individual capacity. The CNA for the pilot countries will serve as an empirical baseline, needed for assessing the degree to which PISA participation builds assessment capacity at any of these levels or categories.

Source: Modified from OECD (2014b), *Capacity Needs Analysis Report: Zambia*, PISA for Development, OECD Publishing, Paris www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf.

Implications

Two main implication of this chapter are that: *i)* greater attention will need to be placed on capacity building activities for PISA; and *ii)* greater effort must be made to empirically evaluate the effects of PISA participation on assessment capacity.

For middle-income countries

To understand the strengths and weaknesses of their own capacity for undertaking international large-scale assessments:

1. Countries should undertake systematic analyses of their assessment capacity. Many countries may not have a good understanding of the scope of assessment operations or of the detailed project management involved. An online tool, such as the CNA, may help countries not only identify the tasks involved in conducting an assessment, but also facilitate the management of these tasks.

For development partners

To enhance capacity building activities:

1. Donors could focus on remediating the assessment weaknesses identified in the CNA's of low- and middle-income countries that are considering participating in PISA. Support (including resources and training) could target these weaknesses as they are identified for each participating country.
2. Donors could also consider supporting or enhancing online self-help tools mapped to the National Project Manager's manual, which could improve the efficiency of management and budgeting related to assessment. For example, the CNA includes a management tool that allows countries to: *i)* co-ordinate their in-country tasks with the international timeline; *ii)* add local tasks; *iii)* allocate resources to and identify dependencies among tasks; and *iv)* incorporate changes in the international timeline into country timelines. This type of tool could be valuable in communicating the complexity of the assessment process itself.

To provide empirical evidence of PISA's effect of assessment capacity:

1. Donors could commission empirical evaluation of the impact of PISA participation on assessment capacity development in middle-income countries. The preparation and design of the PISA for Development initiative offer opportunities to evaluate the impact of participating in PISA on a country's improvement in assessment capacity. Currently, baseline indicators of assessment capacity, which are crucial for assessing change in capacity, have been established through CNAs that have been completed in six of the seven PISA for Development pilot countries (OECD, 2014a, 2014b, 2015a, 2015b, 2015c, 2015 forthcoming). By administering the CNAs after the PISA cycle has been completed it will be possible to identify areas in which participation in PISA has been associated with change and improvement in capacity.[6]

Notes

1. These workshops eventually led to the publications listed in Box 4.1.

2. Esther Duflo, Erik Hanushek, Anil Kanjee, Peter Orazem, Erlinda Pefianco and Paulo Renato Souza.

3. Prior to the SABER-Student Assessment evaluation, four countries had participated in TIMSS only, three countries had participated in PISA only, and three countries had participated in both of these international large-scale assessments.

4. SABER-Student Assessment evaluations for both Lebanon and Viet Nam were completed before they participated in PISA, so these two countries are omitted from this analysis.

5. The fact that around half of the respondent teams reported that their team had not received training is surprising, since all country teams should have received training.

6. A pre-post comparison of assessment capacity can be made, but would not meet standards for impact evaluations.

References

Anderson, P. and **G. Morgan** (2008), *National Assessments of Educational Achievement, Volume 2: Developing Tests and Questionnaires for a National Assessment of Educational Achievement,* The World Bank, Washington DC.

APA Division 5 (2011), Journals in Assessment, Evaluation, Measurement, Psychometrics and Statistics, www.apa.org/divisions/div5/journals.html.

Cartwright, F. (2015), *PISA for Development Planner*, http://polymetrika.org/PISAD.

Clarke, M. (2012), *What Matters Most for Student Assessment Systems: A Framework Paper*, The World Bank, Washington, DC.

Elley, W. (2002), *Evaluating the Impact of TIMSS-R (1999) in Low- and Middle-Income Countries: An Independent Report on the Value of World Bank Support for an International Survey of Achievement in Mathematics and Science*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.

Ferrer, G. (2006), *Educational Assessment Systems in Latin America: Current Practice and Future Challenges*, PREAL, Washington, DC.

Gilmore, A. (2005), *The Impact of PIRLS (2001) and TIMSS (2003) in Low- and Middle-income Countries*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.

Greaney, V. and **T. Kellaghan** (2012), *National Assessments of Educational Achievement, Volume 3: Implementing a National Assessment of Educational Achievement*, The World Bank, Washington, DC.

Greaney, V. and **T. Kellaghan** (2008), *National Assessments of Educational Achievement, Volume 1: Assessing National Achievement Levels in Education,* World Bank Publications, Washington DC.

Kellaghan, T. and **V. Greaney** (2001), *Using Assessment to Improve the Quality of Education*, International Institute for Educational Planning, UNESCO, Paris.

Kellaghan, T., V. Greaney and **S. Murray** (2009), *National Assessments of Educational Achievement, Volume 5: Using the Results of a National Assessment of Educational Achievement*, World Bank Publications, Washington DC.

Large-scale Assessments in Education (2015), www.largescaleassessmentsineducation.com/about.

Lockheed, M. and **S. Cueto** (eds.) (2006), "Education equity and quality in developing and transition countries", [Special Issue] *International Journal of Educational Research*, Vol. 45, No. 3. Elsevier, Amsterdam.

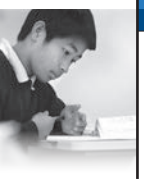
Lockheed, M. (2015), "Why do countries participate in international large-scale assessments? The case of PISA", *Policy Research Working Papers,* WPS 7447, World Bank, Washington DC.

Lockheed, M. (2013), "Causes and consequences of international assessments in developing countries", in H.-D. Meyer and A. Benavot (eds.), *PISA, Power and Policy: The Emergence of Global Educational Governance*, Symposium Books, Didcot, Oxford.

Lockheed, M. (2010), *The Craft of Education Assessment: Does Participating in International and Regional Assessments Build Assessment Capacity in Developing Countries?* International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.

Murimba, S. (2005), "The Impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)", *Prospects,* Vol. 35, No. 1, pp. 91-108. Springer.

Murphy, P., V. Greaney, M. Lockheed and **C. Rohas** (eds.) (1996), *National Assessments: Testing the System*, The World Bank, Washington DC.



OECD (2015a), *Capacity Needs Analysis Report: Guatemala*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Guatemala%20CNA%20reportFINAL_SL.pdf.

OECD (2015b), *Capacity Needs Analysis Report: Ecuador*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Ecuador%20CNA%20ReportFINAL_SL.pdf.

OECD (2015c), *Capacity Needs Analysis Report: Paraguay*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/Paraguay%20CNA%20report_FINAL.pdf.

OECD (forthcoming), *Capacity Needs Analysis Report: Cambodia*, PISA for Development, OECD, Paris.

OECD (2014a), *Capacity Needs Analysis Report: Senegal*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Pisa-D_CNA_Senegal_FINAL_revised.pdf.

OECD (2014b), *Capacity Needs Analysis Report: Zambia*, PISA for Development, OECD, Paris, www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf.

Shiel, G. and **F. Cartwright** (2015), *National Assessments of Educational Achievement, Volume 4: Analyzing Data from a National Assessment of Educational Achievement,* World Bank Publications, Washington, DC.

Wagemaker, H. (2014), "International large-scale assessments: From research to policy", in L. Rutkowski, M.V. Davier, and D. Rutkowski. (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues and Methods of Data Analysis*, Taylor and Francis, New York.

PISA FOR DEVELOPMENT

5

How have PISA results informed education policy discussions and affected policy in middle-income countries?

This chapter explores the extent to which the Programme for International Student Assessment (PISA) has informed education policy in middle-income countries. The first section of the chapter provides a brief review of the education policy process, drawing attention to the important role of discussion and debate in goal-setting and policy formulation. The second section presents new empirical evidence regarding education policy discussions in middle-income countries participating in PISA, drawing from the analysis of media in middle-income case study countries. The third section reviews the evidence of education policy reforms in these countries and their linkages with PISA, drawing on reviews of donor support for assessment, the use of PISA in general policy dialogue, and the use of PISA in identifying specific policy issues related to the quality and equity of education systems. The fourth section examines the evidence related to PISA's impact on policy agendas and country-level reforms. A final section presents conclusions and implications.

INTRODUCTION

The OECD says that PISA "offers insights for education policy and practice" and that its results allow policy makers around the world to "learn from policies and practices applied elsewhere" (OECD, 2014). Research evidence indicates that high-income OECD member countries have responded to PISA results by seeking to learn from the experiences of other countries (Breakspear, 2012; Dixon et al., 2013; Heyneman and Lee, 2014; OECD, 2011). These same sources indicate that PISA has played an increasing role in the educational research literature in high income countries. Less evidence is available for middle-income countries that participate in PISA and it is unclear whether or not PISA has offered insights for education policy and practice in these countries, and if such insights have affected education policy.

This chapter explores these issues in three sections. First, it provides a brief review of the education policy process, drawing attention to the important role of discussion and debate in goal-setting and policy formulation. Second, it presents new evidence regarding education policy discussions in middle-income countries participating in PISA. Third, it reviews the evidence of education policy reforms in these countries and their linkages with PISA. Policy reform is not spontaneously generated, but is the direct outcome of both public discussions (in various media, for example) and private discussions (between ministers of education and international donors, for example). This chapter reviews the evidence regarding these discussions first.

THE EDUCATION POLICY PROCESS

There are many definitions[1] of education policy, but most definitions include two separate features: values and actions. Values inform statements of general intent, while actions are specific procedures. For example, forty years ago, Kogan (1975) observed that policies are "operational statements of values." Two years later, Jennings (1977) presented an influential "linear" model of policy development: evidence of a problem emerges, opinions about the problem crystalise around specific options, policy options are presented formally, further discussion around options occurs, policy makers select the key policy options, and administrative procedures are developed that operationalise the policy (Bell and Stevenson, 2006). The media are particularly important in focusing public opinion around issues, or "agenda setting" (McCombs and Shaw, 1972). The education reform process has been described in terms of agenda setting, reform design, reform adoption, reform implementation, and institutionalisation (Grindle, 2004).

Bell and Stevenson observe that "what is often presented as a policy is frequently no more than a statement of intent, a plan of action or a set of guidelines" and that "policy is about the power to determine what gets done, or not done" (Bell and Stevenson, 2006). In other words, "policy…can be seen as both operational statements of values and as the capacity to operationalise values through the ability to exert influence at key points" in policy development (Bell and Stevenson, 2006: 23).

In recent years, international values in education have been formalised through various declarations and goal statements that have focused on improving access to education and improving the quality of education. For example, the Jomtien Education for All (EFA) goals included achievement of universal access to learning, a focus on equity, and an emphasis on learning outcomes. The Millennium Development Goals (MDGs) for education included goals of achieving universal primary completion and eliminating gender gaps in enrolment at all levels of education. The Sustainable Development Goals (SDG), and associated targets for education, seek to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

Improving learning outcomes – the skills needed for employment in the 21st century[2] – has moved from being one of several goals to being a central education goal of "quality education." PISA and other international large-scale assessments document shortcomings in "quality education" and hence provide empirical evidence of problems to be resolved through public policy. Governments in most countries have considered a wide range of policy options and have implemented actions designed to ameliorate these problems. Some of these policy options are contested and are subject to public debate.

PUBLIC MEDIA DISCUSSIONS OF EDUCATION POLICY

It is unclear how much public debate and discussion of education and education policy takes place in middle-income countries, and how much evidence exists that public discussions of education policy are related to PISA. Public discussions can occur both in media targeted at a wide range of stakeholders, such as television or newspapers, and in publications targeted at academic audiences, such as professional journals. They can also occur in a wide range of non-print media, such as blogs or emails.

Prior research on public discussions has generally focused on print media in high-income OECD countries. For example, research on the press coverage of PISA 2006 showed a sharp increase in the number of articles published in December 2007, following publication of the PISA 2006 results, but this research was limited to four OECD countries: Finland, France, Germany and the United Kingdom (Dixon et al., 2013). Academic literature is much more likely to study education policy and examine the PISA data in detail, but again most research has focused on high-income and OECD countries. For example, among 1590 academic journal articles and book chapters referencing both "PISA" and "education policy" published in the first three quarters of 2015 and accessed through Google Scholar,[3] only 39 (2.5%) also referenced "developing countries."

References to PISA in the media often mention the assessment results but provide no further information. For example, Figazzolo (2009) notes that "out of about 12 000 articles published at worldwide level between December 2007 and October 2008 [mentioning PISA], around 40% make a simple reference to PISA 2006, without further explanation [and] around 30% of the articles make a reference to PISA results in order to advocate for reforms without even analysing what these results mean." Only about 2-3% discussed specific evidence for the suggested reform.

The research reported in this section of this chapter is drawn from Shadrova (2015)[4] and extends these prior analyses by: *i)* focusing on middle-income countries that have participated in PISA; and *ii)* examining references to specific education policy terms, rather than only the term "PISA." This new research is consistent with prior research as it focuses on print media.

The research identifies the changes in the frequency that specific education policy terms appeared in major newspapers in the nine low- and middle-income case study countries (Brazil, Bulgaria, Colombia, Georgia, Indonesia, Jordan, Kyrgyzstan, Turkey, and Viet Nam[5]), following publication of PISA results in the country. The Shadrova (2015) study, commissioned for this chapter, is the first systematic investigation of the public dialogue about PISA and specific policy reforms for low- and middle-income countries. It finds statistically significant relationships between one or two individual educational policy terms and relevant PISA cycles in the publications of Brazil, Bulgaria, Colombia, Turkey and Jordan, but not in Kyrgyzstan. The databases built for Georgia and Viet Nam could not be used due to technical reasons.

Methodology[6]

Publications of record for each of the nine case study countries were identified manually, partly through contacting representatives of the countries involved or experts, and through selecting the largest and oldest archives in each language. The 3 publications with the largest circulation[7] in each country were identified, yielding 27 sources altogether. These publications are listed in Table 5.1; they are generally print publications that have been archived on line.

A list of approximately 75 education policy-related terms based on OECD classifications and the literature regarding education policy was developed and translated into the 9 target languages, using 5 different scripts, and yielding approximately 750 searchable terms. The terms were related to specific policies that could be implemented by governments to improve learning outcomes. In some cases, a term could not be translated into one of the languages.

Table 5.1

Publications searched for discussion on education, by country, 2000-14

Country	Publication title (transliterated as needed)	Dates of archives (varies by publication)	Number of analysable policy terms
Brazil	Folha de S. Paulo, Veja, O Estadao de S. Paulo	2000-2014	10
Bulgaria	Dnevnik, Dneven Trud	2000-2014	9
Colombia	El Tiempo, El Espectador	2000-2014	10
Georgia[1]	Civil.ge, Georgian Times, newsgeorgir.ru	Excluded	0
Indonesia	Koran Tempo, Tempo, Kompas	2000-2014	10
Jordan	Al Ghad, Ad-Dustour, Al Rai	2000-2014	12
Kyrgyzstan	Delo, Radio Azattyk	2008-2014	6
Turkey	Zaman, Sabah, Haberturk, Hurriyet	2000-2014	7
Viet Nam	Five websites	Excluded	0

Note: Table includes only those terms with observed frequency within a publication equal to or greater than 20 in 2014.

1. Georgia was excluded due to an insufficient number of terms that could be found, and Viet Nam was excluded because the standardisation procedure could not be applied.

StatLink http://dx.doi.org/10.1787/888933294059

These terms were grouped into six broad categories; the first four drawn directly from the OECD and the last two reflecting concerns of educators: *i)* selecting and grouping of students; *ii)* resources invested in education (other than teachers); *iii)* resources invested in instructional staff; *iv)* school governance, assessments and accountability; *v)* curriculum and instruction; and *vi)* student learning outcomes.

Each publication was searched for the occurrence of all specified terms using the Google search engine. The frequency of mention of the individual education policy terms in each of the publications was computed for the years 2000 to 2014 (for Kyrgyzstan, only publications 2008-14 were available). Each term's frequency of occurrence was standardised by dividing the frequency of the term by the most frequently occurring term in the particular language for that year and publication (typically words such as "the" or "and"), to produce a "relative frequency" for each term for each year and publication. The last column of Table 5.1 reports the number of education policy terms with relative frequencies sufficient for analysis.

The 19 education policy terms that occurred with sufficient frequency for analysis are identified in Box 5.1. These are (within each group as numbered above): 1) selecting and grouping students (compulsory schooling, mandatory schooling, gifted education, age at school entry, and equality of opportunity); 2) resources invested in education (pre-primary education, length of school day); and 6) student learning outcomes (academic performance, student performance, mathematics achievement, science achievement, reading achievement, mathematics literacy, science literacy, reading literacy, proficiency levels, mathematics proficiency, science proficiency, reading proficiency, competencies). None of the remaining terms that were related to the other three policy categories – 3) resources invested in instructional staff; 4) school governance, assessments and accountability; and 5) curriculum and instruction – occurred with sufficient frequency for analysis. The general terms "students", "teachers", "schools," and "education" were mentioned frequently and were also analysed. Not all terms occurred in every country.

Box 5.1 Education policy terms

1. Selecting and grouping students

- **Compulsory education**
- **Mandatory schooling**
- **Gifted education**
- **Age at school entry**
- **Equality of opportunity**

At-risk students
Grade repetition
Academic tracking
Vocational programmes
Ability grouping
Elite secondary schools

2. Resources invested in education (other than for teachers)

- **Length of school day**
- **Pre-primary education**

Public expenditure on education (budget for education)
Non-salary expenditures on education
Private expenditure on education
School infrastructure
Instructional materials
Textbook availability
Computers/IT availability
Instructional time
Length of school year
Class size
After-school tutoring
After-school lessons

3. Resources invested in instructional staff (not separated in PISA reports)

Teachers' starting salaries
Teacher salary increases
Pre-service teacher training
Requirements to enter the teaching profession
Student-teacher ratio
Teacher professional development
School head professional development (principal, headmaster, etc.)
Female teachers
Teacher unions

4. School governance, assessments and accountability

School autonomy
Appointing teachers (selecting teachers)
Dismissing teachers (firing teachers)
Formulating school budgets
Allocating budget resources within school
School choice
Private schools
Management training
Parental involvement
National assessments
Examinations
Standardised tests
Classroom assessments
Teacher assessments

5. Curriculum and instruction (not in PISA report)

Curriculum reform
Student-centred teaching
Textbook selection
Course offerings
Duration of compulsory schooling
Duration of pre-tertiary schooling
Course content
Opportunity to learn
Inquiry-based learning

6. Student learning

- **Academic performance**
- **Student performance**
- **Mathematics achievement**
- **Science achievement**
- **Reading achievement**
- **Mathematics literacy**
- **Science literacy**
- **Reading literacy**
- **Proficiency levels**
- **Mathematics proficiency**
- **Science proficiency**
- **Reading proficiency**
- **Competencies**

Educational achievement
Skills
Knowledge and understanding
Equality of outcomes
Benchmarks

Note: Bulleted terms in bold occurred frequently enough for analysis in any country.

For each country, three statistical measures were applied. These assessed the relationship between: *i)* the frequency with which the term occurred in the years before, during and after the publication of the country's PISA results, taking into account the different schedules for PISA Plus cycles; *ii)* peaks in term frequencies in the years around the release of PISA results; and *iii)* the correlation among terms within publications. The release of the international results always takes place in December of the year following the collection of data, which is the year referred to in the name of each cycle. For example, data for PISA 2012 was collected throughout 2012 but the international report was not released until December 2013. Statistical differences reported below are $p<.05$ or better.

Shadrova (2015) notes some limitations of the study, including: reliance on Google, inability to compile a complete collection for each website, difficulties in translation terms, the differences among language typologies, and making cross-country comparisons. In addition, online media in some countries could not be accessed for the entire period (2000-2014), and in other countries the results were insufficient for analysis.

Public discussions of education policy in case study countries

Public discussion of education policy in the case study countries did not, in general increase between 2000 and 2014. This is the case for the overall term "education" and for specific education policy terms.

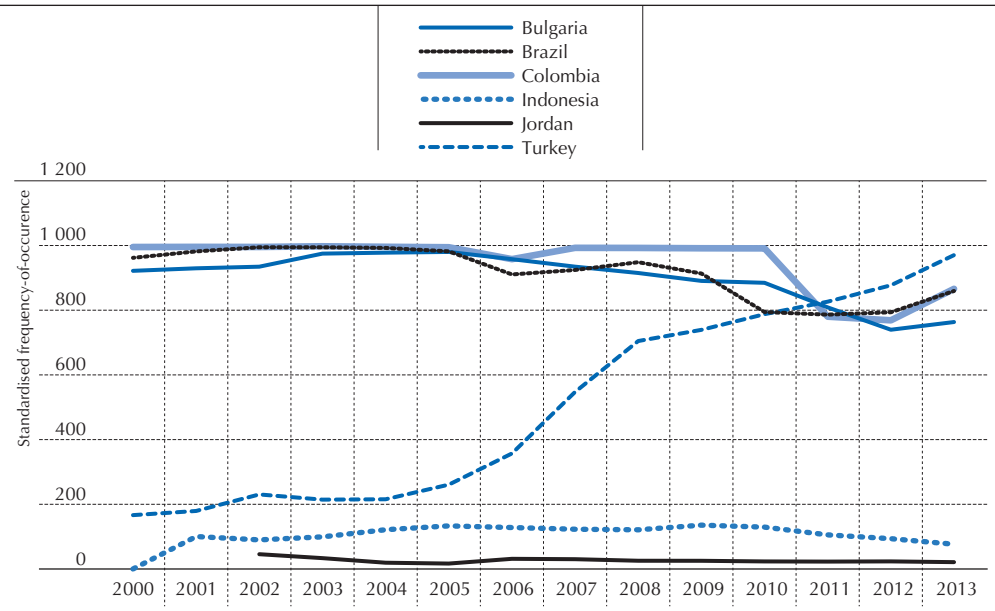
Discussion of "education" increased only in Turkey

The average frequency of occurrence of the term "education" in public media remained relatively stable in five of the six countries from 2000 to 2009, after which it declined slightly. Slight increases occurred after the release of PISA 2012 in Brazil and Colombia. Only one country – Turkey, an OECD member that was undertaking important education reforms beginning in the mid-2000s – showed a steady increase (Figure 5.1). Figure 5.1 presents all six countries on a single scale, but no cross-country comparisons should be made as standardisations are within country only and this figure does not demonstrate greater frequency of term use in Colombia, for example, than in Jordan.

Figure 5.1

Frequency of term "education" in newspapers

6 countries, 2000-14



StatLink http://dx.doi.org/10.1787/888933293898

Discussion of learning outcome terms followed PISA cycle in four countries

Greater attention to student learning outcomes appears related to the PISA cycles. In four of the case study countries (Brazil, Colombia, Jordan and Turkey), statistical tests showed significant relationships between PISA cycles and one or two standardised individual terms, mostly related to student outcomes, in at least one publication. Some terms occurred more frequently during the year following the PISA release compared with other years: Brazil (2 of 10 terms),

Colombia (3 of 10 terms), Jordan (4 of 12 terms) and Turkey (2 of 7 terms). In Kyrgyzstan, the online record was adequate to test only a few relationships, none of which showed any relationship between PISA and the public discourse on education. In Bulgaria only one term in one publication was significantly related to the PISA release. Data from Georgia and Viet Nam could not be analysed, as previously mentioned.

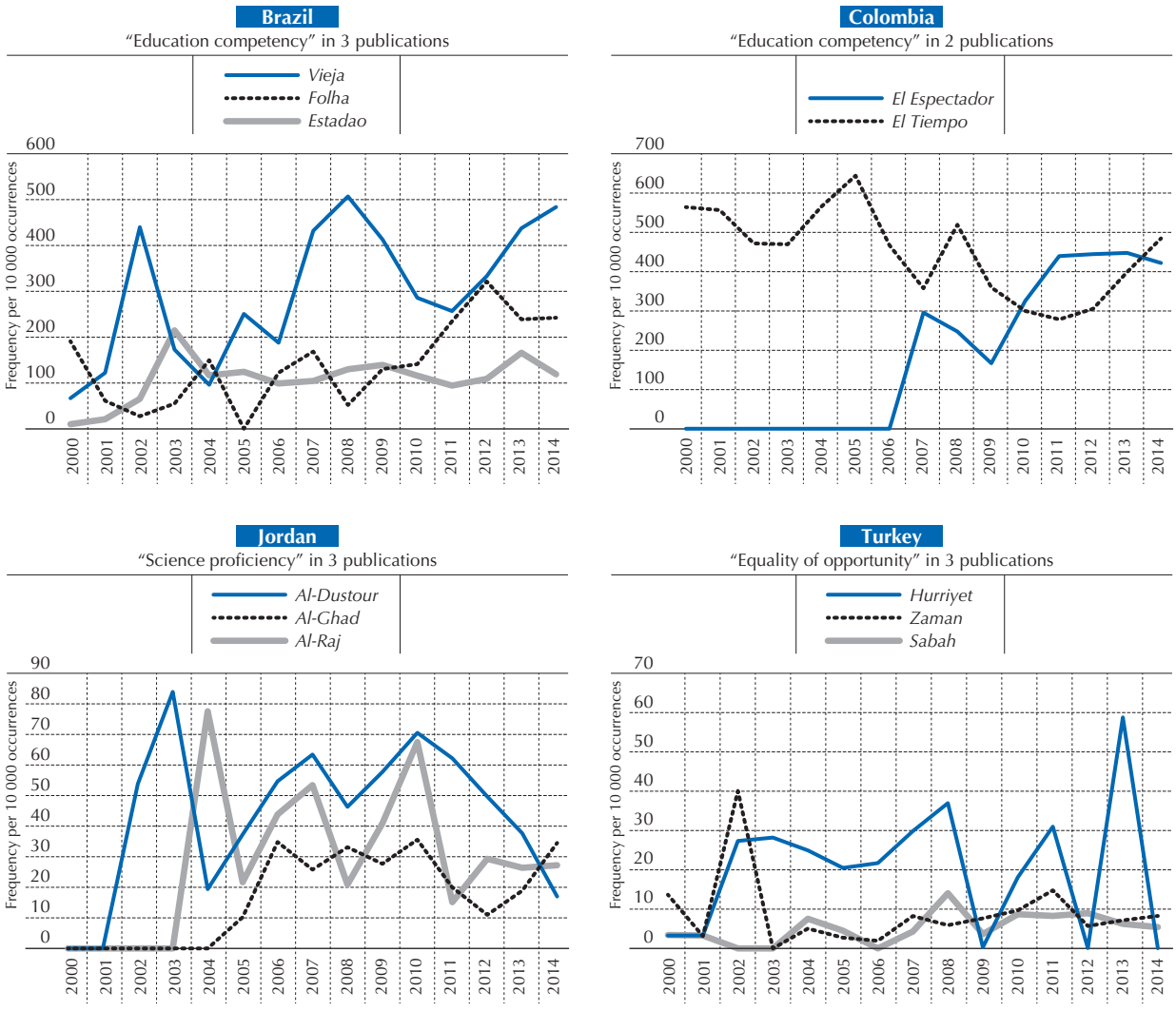
Figure 5.2 provides examples for four of the case study countries of the rise and fall of three specific terms related to student learning outcomes and equity: education competency, science proficiency, and equality of opportunity. Figure 5.2 also demonstrates that results from different publications within the same country were not consistent. In Brazil, Colombia and Jordan, the terms were observed more frequently in the year following the release of the PISA results compared with previous years. This was the case in at least one publication: "education competency" in Brazil (in *Veja* only) and Colombia (in *El Espectador* only), "science proficiency" in Jordan (in *Al Ghad* only), and "equality of opportunity" in Turkey (no statistically significant relationship in any publication).

The figure shows each term's frequency per 10 000 occurrences of the most common word in the various publications. Colombia and Jordan joined PISA for 2006 and Turkey joined for 2003, so observations of terms prior to that time are "non-PISA baseline" levels.

Figure 5.2

Frequency of education terms in publications

Brazil, Colombia, Jordan and Turkey



StatLink http://dx.doi.org/10.1787/888933293898

Discussion of all education policy terms followed PISA cycles in some countries

Another way of looking at whether or not public dialogue about education was influenced by PISA results examined the peaks of all education policy terms taken together and whether these peaks corresponded to PISA cycles in the countries for which suitable data were obtained. No significant relationship between the frequencies of the entire set of policy terms and PISA was found overall, although for some countries a relationship was found for selected years (Shadrova, 2015).

For Colombia and Turkey, the peaks appeared to follow the release of PISA results. In Colombia discussions peaked in 2008, 2011 and 2014 – all approximately one year after the release of the international PISA results for the cycles in which Colombia participated: 2006, 2009 and 2012. Statistically significant relationships between peaks in relative frequency and PISA cycles were observed in *El Espectador*, comparing usage in the year after release of results to usage in other years. In Turkey, peaks were observed in the same years that the international results were released from PISA 2000, PISA 2003, PISA 2009 and PISA 2012. Statistically significant peaks in term usage were observed in *Haberturk*, comparing usage in the study year with usage in other years. Given the steady rise in discussions of education in Turkey over this time period, however, conclusions should not link this growth to PISA.

In Jordan, discussions were very active from 2003, but this may be related to the Trends in International Mathematics and Science Study (TIMSS) 2003, in which Jordan participated, rather than to PISA 2003, in which Jordan did not participate. Subsequent peaks in Jordan were observed in 2009, the year of the main study for PISA 2009, and in 2013, the year results from PISA 2012 were released.

Public discussion of education topics was also observed in Indonesia, which has a long history of participation in international assessments. Peaks in education terms were highest for 2004, the year the results of both PISA 2003 and TIMSS 2003 were released, and in 2012, the year that the results of TIMSS 2011 were released, but the differences were not statistically significant. In Indonesia and Brazil, significant differences were found for years that were entirely unrelated to the PISA cycles.[8]

Discussions of specific education policy terms did not generally follow PISA cycles

Public discussion of most of the individual education policy terms did not increase during the year immediately following the release of PISA in the case study countries compared with previous years. Across all countries, publications and PISA cycles, the years in which PISA results were released in the countries were not associated with increases in the frequency with which terms were mentioned in newspapers. While some terms rose and fell in three-year patterns that matched either the years that PISA results were released for each country or the years thereafter, the same was not true for the majority of terms. In addition, the term "OECD" did not occur more frequently the year after the PISA results were released compared to previous years in any of the countries. This is in sharp contrast to other studies that found an increased visibility of "OECD's PISA" during the year following the release of PISA results in OECD countries.

No coherent clusters of policy terms

Another way of observing a coherent level of discourse about education policy was through the clustering together of education policy terms within publications. However, repeating clusters of term frequency correlations could not be found for any of the countries, suggesting individual differences by both country and publication rather than a fixed set of topics publicly discussed. In some publications in three countries there was an association between the term "OECD" and one or more policy terms: *i)* in Brazil, the terms "mathematics competency", "equality of opportunity" and "preschool" had correlations greater than 0.40 with the term "OECD"; *ii)* in Colombia, the terms "education competence" and "mathematics competence" were correlated with the term "OECD"; and *iii)* in Turkey, the term "OECD" was correlated with all the following terms: "preschool", "education competence", "equality of opportunity" and "literacy" (Shadrova, 2015). With the exception of "preschool", no specific policy term related to a policy reform was correlated with any other term.

Reports from countries emphasise PISA results, not policy options

For evidence from research to inform public dialogue about education policy it is necessary for the results to be widely disseminated, often through mass media such as newspapers, television, radio and online resources. Theoretical and empirical work indicate that "access to mass media empowers people politically" (Olper and Swinnen, 2013; Stromberg, 2001). In OECD countries, a great deal of dissemination of research related to education policy has taken place over the past 15 years, often peaking with the release of PISA or TIMSS. For the middle-income case study countries, however, direct evidence from the reviewed publications indicates substantially less public discussion of education policy topics.

Country management teams for PISA 2015, who were interviewed in late 2014, reported that dissemination of PISA results had occurred. The national project managers for Bulgaria, Colombia and Viet Nam recalled that editorials and front-page articles had covered PISA results, and national teams from Brazil, Bulgaria, Colombia and Georgia all noted that PISA results were discussed in online blogs. They confirm, as is the case in many OECD countries, that public discussions focus on student learning outcomes (that is, PISA results) rather than on specific education policies that may be related to improvements in outcomes.

The next section of this chapter explores how PISA and other international large-scale assessments have influenced private dialogue between governments and international donors in the case study countries. International donors are singled out for this section because they play a large role in agenda setting for reform and because their dialogue is publicly available (Parks et al., 2015). This dialogue is typically focused on which education policies a country could consider adopting to improve learning outcomes.

PRIVATE DISCUSSIONS AND DIALOGUE BETWEEN GOVERNMENTS AND INTERNATIONAL DONORS

International large-scale assessments are used in three types of dialogue between donors and client countries: in general policy documents that do not target specific countries; in the identification of issues to be addressed through projects and programmes that often support policy reforms in specific countries; and in monitoring project and programme outcomes in these countries. Possibly as a consequence of this dialogue, donors have also provided financial support that has enabled countries to participate in international large-scale assessments. This section reviews donor support for assessment and the uses of assessment, and focuses on the experience of the World Bank, which is recognised as the largest development partner in education. It draws on several sources: *i)* previous reviews of World Bank projects, 1975-1993 and 1998-2009; *ii)* a review of World Bank project appraisal documents in education, 2009-2014, specifically undertaken for this report; and *iii)* World Bank implementation completion reports for selected earlier projects.

Donor support for assessment

The World Bank has supported various types of student assessments since 1975. A review of Bank support for "testing" 1975-1992 found 85 projects with testing subcomponents, most of which supported examinations and national assessments (Larach and Lockheed, 1992). The share of projects with testing components increased sharply between the late 1970s and the early 1990s, from fewer than 15% to over 40% of education projects. International large-scale assessments were not singled out in this review.

Table 5.2

Recent World Bank projects mentioning international large-scale assessments

Name of project	Year project approved	Rationale for country's project	Project's development objective	Project may support country's participation in international large-scale assessment
Kyrgyzstan: Sector Support for Education Reform Project	2013	PISA 2006 and 2009 identified low achievement, geographical inequities	National assessment	No
Senegal: Quality and Equity of Basic Education	2013	PASEC identified low achievement	No	PIRLS, PISA, TIMSS
Peru: Basic Education Project	2012	No	No	PISA, TERCE
Moldova: Education Reform Project	2012	PISA 2009 + identified low performance	Participation in PISA 2015	PISA
Sri Lanka: Transforming the School Education System as the Foundation of a Knowledge Hub Project	2011	No	National assessment	PISA, TIMSS
Nicaragua: Second Support to the Education Sector Project	2011	SERCE identified low achievement and regional inequities	National assessment	TERCE
Lebanon: Second Education Development Project	2010	Yes	No	PISA, TIMSS
Jordan: Second Education Reform for the Knowledge Economy Project	2009	Yes	National assessment	PISA, TIMSS

Note: Financial resources from the project may be used for a country's participation in an international large-scale assessment; the country may or may not actually participate in any specific international large-scale assessment.
Source: World Bank (2015a), Projects database, http://go.worldbank.org/KTPE1WKU20.

A more recent review of 166 World Bank projects approved from 1998-2009 concluded that over 75% of projects supported some type of assessment (Liberman and Clarke, 2012). Approximately one third of these projects supported one or more international large-scale assessment: 8 projects supported the Progress in International Reading Literacy Study (PIRLS), 19 projects supported PISA, and 21 projects supported TIMSS.

A search of the World Bank's electronic database of project documents, conducted for this report, found continuing support for assessments. A slightly higher percentage of the 18 education projects approved from 2010 to 2014 supported international large-scale assessments (38%), compared with projects approved from 1998-2009 (33%) Among these 18 education projects, 7 projects supported the country's participation in one or more international large-scale assessment (these are listed in Table 5.2). Only one project specifically indicated the level of support for PISA: approximately USD 2 million was allocated for Peru's participation in PISA 2015.

General policy dialogue with development partners

The World Bank has utilised the results of international large-scale assessments in its general policy dialogue with countries for over 35 years. In 1978, economists John Simmons, then at the World Bank, and Leigh Alexander published a literature review of the determinants of school achievement in developing countries that included the early studies of reading, science and mathematics carried out by the International Association for the Evaluation of Educational Achievement (IEA) (Simmons and Alexander, 1978). This research was cited in the first World Bank education sector policy paper of 1980 (World Bank, 1980).

Some years later, Simmons' review was expanded upon by Bruce Fuller (Fuller, 1987; Fuller and Clarke, 1994), then a sociologist at the World Bank, and the results were included in several World Bank policy documents in the early 1990s. These documents included: the Bank's contribution to the 1990 World Conference on Education For All (the Jomtien conference), the widely cited *Improving Primary Education in Developing Countries* (Lockheed and Verspoor, 1990), and the Bank's primary education policy paper of 1990 (World Bank, 1990).

World Bank education policy papers published over the past 20 years have continued to draw on the results from international large-scale assessments, including PISA. The 1995 *Priorities and Strategies for Education* cites the results of TIMSS and previous IEA studies and played a key role in increasing the focus on measuring and monitoring learning outcomes in countries (World Bank, 1995). The *Education Sector Strategy Paper* of 1999 encourages countries to: "1) establish standards for what students should know and be able to do at various stages of the education system; 2) participate in international evaluations of educational achievement; and 3) develop good national assessment systems." (World Bank, 1999). The World Bank Education Strategy 2020 (World Bank, 2011) emphasises the importance of measuring learning outcomes through national, regional and international assessments.

Identification of policy issues: Quality and equity

Results from some international large-scale assessments informed many of the World Bank's education projects of the 1990s, before the establishment of PISA and when achieving universal primary education was still a challenge for many of the Bank's client countries. The international large-scale assessment results provided some of the first empirical evidence for the relative effectiveness and efficiency of specific education investment (Lockheed and Hanushek, 1988). Results demonstrating the importance of learning materials and teacher preparation were influential in raising the number of World Bank projects supporting textbook development and teacher in-service training throughout the 1990s.

PISA and other international large-scale assessments have continued to be analysed to support the need for education system improvement in low- and middle-income countries. Table 5.2 summarises how PISA and other international large-scale assessments have been used in the project appraisal documents of recent World Bank projects to identify the need for financial support from the World Bank for the education sector in the country and to measure project outcomes, increasingly important with the rise of results-based-funding.

In addition to the World Bank; the Inter-American Development Bank, the Millennium Challenge Corporation and the Asia Development Bank have used PISA results to identify not only a country's overall performance relative to OECD averages, but also to locate between and within school inequities. For example, PISA results regarding low performance and/or sharp urban-rural differences in student learning were incorporated into rationales for projects supported by the Millennium Challenge Corporation in Georgia and the World Bank in Bulgaria,[9] the Republic of Moldova and Kyrgyzstan.[10] Other international large-scale assessments – principally the *Programme d'Analyse des Systèmes Educatif de CONFEMEN* (PASEC) and the Regional Comparative and Explanatory Study (ERCE) – have also been used for this purpose.

Project and programme results-monitoring and conditionality

World Bank projects and programmes, particularly in the early-to-mid 2000s, have utilised international large-scale assessments for both results-monitoring and conditionality. For example, Jordan's first Education Reform for the Knowledge Economy project (ERfKE I, 2003-2008) used TIMSS as an outcome measure, with the end target objective for the project "a positive statistically significant improvement in 2007 TIMSS scores." Although ERfKE II (2009-15) could have adopted PISA or TIMSS as an outcome indicator, a decision was made to use a national assessment tool instead. In one unusual case, Bulgaria's second Social Sectors Institutional Reform development policy loan from the World Bank required the successful completion of PISA 2006, PIRLS 2006 and TIMSS 2007 as a condition for the loan.

Overall, international large-scale assessments have contributed to the understanding of education policy issues in low- and middle-income countries for several decades. These assessments have informed dialogue about education quality and equity and have helped identify promising pathways for education reform. Results for specific countries have helped shaped investment priorities.

IMPACT OF INTERNATIONAL ASSESSMENTS ON POLICY AGENDAS AND REFORMS

The impact of an international assessment such as PISA on education policy agendas and reforms in any country can be direct or indirect. Direct effects could come from discussions among stakeholders at the country level, whereas indirect effects could come from general discussions in broader policy forums, conferences or the media. International large-scale assessments could affect agenda setting, policy formulation, policy implementation, or monitoring and evaluation. Impacts could also include the confirmation or rejection of existing education policies or policy directions. Perhaps the most important impact is the identification of education issues.

PISA and other international large-scale assessments have often drawn attention to three pressing issues in education: *i)* the poor quality of student learning; *ii)* inequality in student learning; and *iii)* inequality in the distribution of learning resources across schools and geographical areas (OECD, 2013a, 2013b). Interviews carried out for this report confirmed that PISA draws attention to these issues. In Turkey, for example, one interviewee noted that "Turkey engaged with PISA as a key component and source of evidence for the preparation of the Education Sector Study programme in 2004-05", and that "the government agreed to join PISA, work with national and international experts to analyse PISA results to learn about education quality and equity in Turkey, and publicise the results broadly to citizens to raise education quality as a key issue." In Colombia, one interviewee mentioned that the World Bank had "a very deep discussion" with the minister of education about the results of PISA and the minister was interested in learning more about "some international experiences [from] Poland, Finland, Canada and Brazil."

Methodological caveat

Empirically linking PISA with policy reforms in middle-income countries is a difficult task for four methodological reasons. First, the PISA cycles occur frequently and concurrently with many other events that could contribute to reforms, including changes in government, so that empirical attribution of reforms to PISA, specifically, is virtually impossible. Second, as the previous sections have demonstrated, the public discussions of PISA results are not always visible enough to stimulate reforms in these countries and do not appear linked to the PISA cycle. Third, most empirical studies have examined only the perceived, rather than the actual, effects of PISA on education policy and therefore shed only a weak light on PISA effects. Finally, even the surveys of perceptions about PISA have focused mainly on OECD and high-income countries, with middle-income countries less well represented. Some of these caveats also apply to research on the policy effects of other international large-scale assessments.

Research on international large-scale assessments and education policy

Research on education policies can be divided into two main groups: research related to the stages of policy formation and implementation, and research related to specific education policies. PISA and other international large-scale assessments often contribute to one important stage in policy formation – agenda setting – through documenting issues of education quality and equity. Analyses of the results from international large-scale assessments also often suggest areas where policy reforms could improve these outcomes. Some research suggests that international large-scale assessments influence both policy agenda setting and policy implementation.

International large-scale assessment may affect policy agenda setting

Two studies that examine the impact of international large-scale assessments on education policy agenda setting reach different conclusions. The first, a systematic review of the literature from 1990-2011, identified 19 studies that addressed the impact of international assessments on implemented education policies in low- and middle-income countries;

of these, 11 were considered "high quality" studies, and three of the "high quality" studies referenced PISA, specifically (Best et al., 2013a). Best's review grouped the effects of PISA with those of other international large-scale assessments (TIMSS, PIRLS, International Assessment of Education Progress [IAEP], IEA Civic Education Study (CIVED), International Computer Competence Study [ICCS], Monitoring Learning Achievements [MLA]) to draw conclusions regarding assessment effects on four policy processes: agenda setting, policy formulation, policy implementation, and monitoring and evaluation. The review found that international large-scale assessments were mainly associated with policy implementation and monitoring and evaluation, and less with agenda setting or policy formation.

The second study, a recent survey of 6 744 "opinion leaders"[11] in 126 low- and middle-income countries and jurisdictions,[12] examined the policy-making influence of external assessments across multiple sectors, including education. The study concluded that external assessments, including PISA, were more influential on agenda setting than on specific policy design.[13] Actual education reforms were perceived as only weakly influenced by external partners, including the OECD (Parks et al., 2015).[14]

International large-scale assessments may suggest specific education policies

International large-scale assessments are perceived to influence a variety of specific education policies that are designed to improve education quality and equity in middle-income countries.

Recent reviews by the OECD (OECD, 2013b) and others (Glewwe et al., 2014; Kremer, Brannen and Glennerster, 2013; Krishnaratne et al., 2013; McEwan, 2014; Murnane and Ganimian, 2014), suggest five broad areas where changes in education policy could result in higher student learning outcomes in low- and middle-income countries: *i)* selecting and grouping students; *ii)* non-personnel resources invested in education; *iii)* resources invested in the quality of instructional staff; *iv)* school governance and assessments; and *v)* curriculum and instruction.

Within each of these categories are numerous specific policy areas, for example:

1. Selecting and grouping students: class size (Heyneman, 2003; Lockheed and Hanushek, 1988) or grade retention (Manacorda, 2012; OECD, 2013a; Schwerdt and West, 2012).
2. Resources invested in education: instructional materials and technology (Baker, Goesling and LeTendre, 2002; Lockheed and Hanushek, 1988; Lockheed and Verspoor, 1990; OECD, 2013b).
3. Resources invested in the quality of instructional staff, particularly in-service professional development (Abdul-Hamid, Abu-Lebdeh and Patrinos, 2011; Best et al., 2013; Glewwe et al., 2014; Hanushek, Link and Wößmann, 2013; Wagemaker, 2014).
4. School governance, accountability, assessment and standards (Best et al., 2013; Galiani and Perez-Truglia, 2014; Hanushek and Wößmann, 2014).
5. Curriculum and instruction, including pedagogy, teaching and opportunity to learn (Best et al., 2013; Heyneman and Lee, 2014; Schmidt, Zoido and Cogan, 2014; Wagemaker, 2014).

Research on PISA and education policy

Several studies have examined PISA's effect on education policy, but most have reported results largely from high-income countries. A recent review of education policy in Latin America provides extensive descriptions of policy changes, but makes little connection with PISA (Rivas, 2015). Studies of PISA effects an education policy include:

1. A very comprehensive OECD survey of PISA stakeholders examined the policy effects of PISA 2000, 2003 and 2006. It achieved a very high response rate from the surveyed high-income countries (86%, 31 out of 36 countries), but a much lower response rate from the surveyed middle-income countries (57%, 12 out of 21 countries) (OECD, 2008).
2. Only 5 of the 37 PISA Governing Board members who responded to a 2011 survey of all members represented middle-income countries (Breakspear, 2012).
3. A recent study of PISA's effect on education policy included no middle-income countries, but focused on four European Union (EU) countries participating in PISA 2006 and PIRLS 2006 (Dixon et al., 2013).
4. An OECD study summarised policies and practices from eight high-income economies and two upper-middle-income countries that could be applicable for the United States (OECD, 2011).
5. An Education International (EI) study surveyed EI representatives from 26 countries, of which 23 were high-income countries (Figazzolo, 2009).

6. A narrative review provided examples of the effects of assessments in 21 countries, including 5 middle-income countries. PISA was mentioned as a driver of change in Mexico (Heyneman and Lee, 2014).
7. A study of the policy effects of PISA focused on six high-income countries and economies (Baird et al., 2011).
8. A study of the "soft policy" effects of PISA focused on Switzerland and the United States (Bieber and Martens, 2011).

The absence of data from most middle-income countries means that the conclusions reached from many prior studies of the effects of PISA on education policy may be less applicable to middle-income countries than to the high-income OECD countries for which the studies were intended.

Differences between middle-income countries and high-income countries

The perceptions of stakeholders from middle-income countries regarding the effects of PISA on education policy differ from the perceptions of stakeholders from high-income OECD countries. An OECD survey of stakeholders from countries participating in the first three PISA cycles reported that PISA positively affected education policy with respect to: development of national standards, establishment of national institutes of evaluation, changes in the curriculum, introduction of targeted educational programmes, increased allocation of resources to schools and increased collaboration among key stakeholders (OECD, 2008). However, among the 12 middle-income countries that responded to this survey, 7 reported that PISA had "relatively low levels of impact on policy formation", 2 reported "relatively medium levels of impact" and only 3 (Kyrgyzstan, Mexico and Thailand) reported "relatively high levels of impact" (OECD, 2008).

Similar differences were found in a survey conducted in 2013 (Breakspear, 2012). Nearly 60% of the 32 high-income countries/economies covered reported that PISA had been "extremely" or "very" influential in informing the policy-making process. However, middle-income countries were somewhat less positive, and the small number of respondents meant that these differences could not be tested. Among the five middle-income countries, none reported that PISA had been "extremely" influential and only Hungary and Mexico reported that PISA had been "very" influential in informing the policy-making process (40% of the middle-income respondents). By comparison, Chile reported that PISA was "moderately" influential and Indonesia and Turkey reported it was "not very" influential, although respondents noted that PISA was used for PISA-referenced performance targets and indicators in both countries.

Moreover, reforms suggested by analyses of PISA in OECD countries – such as increasing school autonomy – appear to be correlated with quality improvements in high-income countries, but not in low- or middle-income countries (Hanushek et al., 2013). As a recent World Bank analysis of data from Bulgaria observes: "the effects of Bulgaria's 2007 school autonomy reform on student achievement are mixed and worse than expected" (World Bank, 2012). Similar conclusions were drawn about the effects of Kyrgyzstan's teacher incentive reform on enhancing teacher motivation (Lockheed, 2014).

Effects of PISA on education policy in middle-income countries

Studies focused on specific middle-income countries, as well as interviews conducted for this report, indicate that the effects of PISA typically relate to revising curriculum standards, establishing performance targets related to PISA, and – in some cases – specific education reform policies intended to boost performance.

For example, Jordan responded to both TIMSS and PISA results to compare itself with the world's best achievers, review its curriculum, establish performance benchmarks and revise teacher training (Abdul-Hamid et al., 2011). In Kyrgyzstan, PISA results affected reforms such as the development of new standards and curricula, reductions in teaching load, upgrading of physical facilities, teaching practices and per-capita financing; although some of these reforms pre-dated PISA (Shamatov, 2014; Shamatov and Sainazarov, 2010).

In addition, middle-income countries have reflected on the types of skills assessed by PISA and how these skills could be better included in the national curriculum and standards. Interviewees from middle-income countries participating in PISA 2015 generally agreed that the results of PISA have informed curriculum improvement and teacher training programmes. In Indonesia, one interviewee noted that the 2013 curriculum "came out of nowhere" but that a "big motivation for the reform was that Indonesia did not do well on PISA"; a new curriculum was developed to respond to "PISA-like things." In Kazakhstan, one interviewee noted that PISA had been used with respect to "the National Action Plan on Development of Functional Literacy of School Students for 2012-2016 [which] was adopted on the instruction of the President [in his] annual Address to the Nation." In Mexico, the president's education sector programme established performance targets for 2012 based on PISA (OECD, 2010). Subsequently, Breakspear (2012: 26-27) reports that for Mexico "The Agreement for the Articulation of Basic Education asserts that the whole curriculum should set a vision for 2021 that includes generalising the competencies described in PISA Level 3." In Turkey, Breakspear notes that: "national performance targets are determined according to the country's score on PISA and the OECD average."

The experience of PISA has influenced the design of national assessments. Interviewees from middle-income countries participating in PISA 2015 noted that PISA had informed other assessment activities in their country, such as national assessments. For example, in Indonesia the national assessment was being revised to include new PISA-type items. In Moldova, one interviewee observed that PISA results influenced a change in the methodology of the national examination toward greater standardisation. The movement towards the greater use of performance assessment was also noted.

CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES

Conclusions

This chapter has explored how PISA could affect education policy in middle-income countries. The evidence indicates that high-income OECD member countries have responded to the publication of PISA results by seeking to learn from the experience of other countries and by reflecting on education policy, often very publicly. The evidence for middle-income countries is more limited, but suggests that this is less the case.

Public dialogue about education policy, as observed in selected public media from eight middle-income countries, shows little relationship with the publication of PISA results. However, PISA and other international large-scale assessments have informed the private policy dialogue between countries and development partners, as indicated by a review of World Bank projects. International large-scale assessments have provided empirical evidence of the need for policy reforms to improve the quality and equity of education outcomes in numerous countries, and have occasionally been used as key results indicators for projects and programmes supported by donors. This chapter reaches six conclusions:

1. Media coverage of education and education policy does not rise and fall systematically with the publication of PISA results, suggesting that the media may not play as strong a role in education policy agenda setting in middle-income countries as it does in high-income countries and economies.
2. PISA provides evidence of education quality and equity in middle-income countries, and countries have used this evidence in education policy dialogue with development partners.
3. Recommendations for policy reforms, derived largely from analysis of PISA results from high-income countries, may not apply to middle-income countries.
4. PISA's perceived influence on education policy is more positive in high-income countries or economies than in middle-income countries.
5. There is conflicting evidence with respect to the impact of PISA on education policy agenda setting in low- and middle-income countries.
6. Some evidence indicates that the principal education policy reforms influenced by PISA involve curriculum alignment with international standards, teacher training aligned with the curriculum, and improvements in the overall student assessment systems.

Implications

Two main implications to increase the influence of PISA on education policy in middle-income countries can be drawn from these findings.

For middle-income countries

To increase public discussion of PISA results and education policy, in general:

1. Countries could consider expanding media coverage for disseminating the results of PISA beyond conventional press releases that appear to be short-lived in middle-income countries. Approaches could include: using social media to communicate results in local languages and hosting seminars to disseminate information to ministry staff, local NGOs, and local donors.
2. Countries could also expand their outreach to all stakeholders in the education system, including parents, teachers and schools.

For the OECD

To improve the relevance of PISA results for middle-income countries:

1. The OECD could consider revising some of its regular analyses of PISA to focus on middle-income country participants, rather than grouping all non-OECD partner countries/economies into a single group. Instead it could analyse results separately by income group and region.

Notes

1. The OECD defines policy as any course of action undertaken by a government.

2. The term "21st century skills" refers to a broad set of knowledge, skills, work habits and character traits that are believed – by educators, school reformers, college professors, employers and others to be critically important to success in today's world, particularly in collegiate programmes and contemporary careers and workplaces.

3. Google Scholar search on 17 September, 2015 for articles and book chapters containing the words "Programme for International Student Assessment", "developing country" and "education policy".

4. The full report is available from the OECD.

5. These countries participated in PISA in the following cycles: Georgia 2009; Indonesia 2000-12; Viet Nam 2012; Kyrgyzstan 2006, 2009; Bulgaria 2000, 2006-12; Brazil 2000-12; Colombia 2006-12; Jordan 2009-12; Turkey 2003-12.

6. Complete information on the research methodology is provided in Shadrova (2015).

7. The publications with the largest daily (unless otherwise noted) circulation were: Brazil (Veha: 1.2 million weekly), Bulgaria (Dneven Trud: 70 000) Colombia (El Tiempo: 1.3 million daily), Georgia (civil.ge: 10 000 visitors), Indonesia (Kompas: 5 million), Jordan (Al Rai: 80 000), Kyrgyzstan (Vecherniy Bishkek: 50 000).

8. In Brazil, $p = 0.03$ for peaks of 2002, 2007, 2011, 2014; in Indonesia $p = 0.025$ for peaks of 2004, 2009, 2014.

9. Bulgaria DPL I (2007) supported four testing activities, including "managing Bulgaria's participation in international achievement tests such as TIMSS, PISA and PIRLS" (World Bank PAD). The implementation completion report (ICR) for this loan noted that: "In 2006, Bulgaria participated in PISA and PIRLS, and in TIMSS in 2007. PISA results suggest a large variance in student achievement between schools, as opposed to within schools, suggesting that there are many high quality but also many failing schools often in socially excluded Roma localities."

10. Kyrgyzstan (P113350) project appraisal document noted that: "The country's participation in the OECD's Programme for International Student Assessment (PISA) showed that, while Kyrgyz student's performance improved between 2006 and 2009, 15-year-olds lag an approximate four and a half grade levels behind the OECD average."

11. Only 5% of the respondents (377) self-identified as working in the education sector, and only 15 respondents rated PISA's influence.

12. The survey included respondents from internationally unrecognised jurisdictions, such as "Kurdistan" and "Puntland" and therefore the term "country" is not used.

13. PISA was singled out in this report as the assessment having the greatest influence on education policy agenda setting, but no other international large-scale assessment was included in the list of 103 external assessments presented to the respondents, for comparison. Among the surveyed countries, eighteen were participating in PISA 2015. Respondents were asked to identify the most influential assessment specific to their sector, in terms of agenda setting and reform design. Although approximately 5% of the respondents were education experts, not all were asked all questions; 17 respondents – half from the European and Central Asia region – rated PISA positively for its effect on agenda setting for education and 15 indicated that PISA had a positive effect on reform design.

14. Only 15 respondents rated PISA's influence, and no other international large-scale assessment was included in the list of potential influencers. The only other listed development partners in education were: the UNESCO Global Monitoring Report, the World Bank's Education Sector Review, the World Bank's EdStats and the Paris Declaration indicators. The exclusion of the other major international large-scale assessments (TIMSS, PIRLS, SACMEQ, TERCE) from the list raises questions about the comprehensive nature of this study, with respect to education.

References

Abdul-Hamid, H., K. Abu-Lebdeh and **H. Patrinos** (2011), "Assessment Testing Can Be Used to Inform Policy Decisions", *Policy Research Working Papers,* WPS5890, World Bank, Washington, DC., http://dx.doi.org/10.1596/1813-9450-5890.

Baird, J., T. Isaacs, S. Johnson, G. Stobart, G. Yu, T. Sprague and **R. Daugherty** (2011), *Policy Effects of PISA*, Oxford University Centre for Educational Assessment, Oxford, UK.

Baker, D., B. Goesling and **G. LeTendre** (2002), "Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the "Heyneman-Loxley Effect" on Mathematics and Science Achievement", *Comparative Education Review*, Vol. 46, No. 3, pp. 291-312, Univerity of Chicago Press, Chicago.

Bell, L. and **H Stevenson** (2006), *Education Policy: Process,Themes and Impact,* Routledge.

Best, M., P. Knight, P. Lietz, C. Lockwood, D. Nugroho and **M. Tobin** (2013), *The Impact of National and International Assessment Programmes on Education Policy, particularly Policies regarding Resource Allocation and Teaching and Learning Practices in Developing Countries*, final report, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.

Bieber, T. and **K. Martens** (2011), "The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US", *European Journal of Education*, Vol. 46, No. 1, pp. 101-116, John Wiley and Sons, Hoboken.

Breakspear, S. (2012), "The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance", *OECD Education Working Papers*, No. 71, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k9fdfqffr28-en.

Dixon, R., C. Arndt, M. Mullers, J. Vakkuri, K. Engblom-Pelkkala and **C. Hood** (2013), "A Lever for Improvement or a Magnet for Blame? Press and Political Responses to International Educational Rankings in Four EU Countries", *Public Administration*, Vol. 91, No. 2, pp. 484-505, John Wiley and Sons, Hoboken.

Figazzolo, L. (2009), *Impact of PISA 2006 on the Education Policy Debate*, Education International.

Fuller, B. (1987), "What school factors raise achievement in the third world?", *Review of Educational Research*, Vol. 57, No.3, pp. 255-292, Sage, New York.

Fuller, B. and **P. Clarke** (1994), "Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy", *Review of Educational Research*, Vol. 64, No.1, pp. 119-157, Sage, New York.

Galiani, S. and **R. Perez-Truglia** (2014), "School management in developing countries", in P. Glewwe (ed.), *Education Policy in Developing Countries*, University of Chicago Press, Chicago.

Glewwe, P., E. Hanushek, S. Humpage and **R. Ravia** (2014), "School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010", in P. Glewwe (ed.), *Educational Policy in Developing Countries*, University of Chicago Press, Chicago.

Grindle, M. (2004), *Despite the Odds: The Contentious Politics of Education Reform*, Princeton University Press, Princeton.

Hanushek, E. and **L. Wößmann** (2014), "Institutional structures of the education system and student achievement: a review of cross-country economic research", in R. Strietholt, W. Bos, J. Gustafsson, Rosen and M. Munster (eds.), *Educational Policy Evaluation through International Comparative Assessments*, pp. 145-175, Waxman, Germany.

Hanushek, E., S. Link and **L. Wößmann** (2013), "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA", *Journal of Development Economics,* Vol. 104, pp. 212-232, Elsevier, Amsterdam.

Heyneman, S. (2003), "The history and problems in the making of education policy at the World Bank 1960-2000", *International Journal of Educational Development*, Vol. 23, No. 3, pp. 315-337, Elsivier, Amsterdam.

Heyneman, S. and **B. Lee** (2014), "Impact of international studies of academic achievement on policy and research", in L. Rutkowski, M. von Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment,* Taylor and Francis, New York.

Jennings, R. (1977), *Education and Politics: Policy Making in Local Education Authorities*, Batsford, London.

Kogan, M. (1975), *Educational Policy Making: A Study of Interest Groups and Parliament*. Allen and Unwin, London

Kremer, M., C. Brannen and **R. Glennerster** (2013), "The Challenge of Education and Learning in the Developing World" in *Science*, Vol. 340, No. 6130, pp. 297-300, American Association for the Advancement of Science, Washington, DC.

Krishnaratne, S., H. White and **E. Carpenter** (2013), *Quality Education forAallCchildren? What Works inEducation in Developing Countries*, Working Paper 20, International Initiative for Impact Evaluation, New Delhi.

Larach, L. and **M. Lockheed** (1992), *World Bank Support for Capacity Building: The Challenge of Educational Assessment*, World Bank, Washington, DC.

Liberman, J. and **M. Clarke** (2012), *A Review of World Bank Support for Student Assessment Activities in Client Countries, 1998-2009*, World Bank, Washington DC.

Lockheed, M. (2014), "Teacher Opinions on Performance Incentives: Evidence from the Kyrgyz Republic", *Policy Research Working Papers,* WPS 6752, World Bank, Washington, DC.

Lockheed, M. and **A. Verspoor** (1990), *Improving Primary Education in Developing Countries*, Oxford University Press, New York.

Lockheed, M. and **E. Hanushek** (1988), "Improving Educational Efficiency in Developing Countries: What do we Know?" *Compare*, Vol. 18, No.1, pp. 21-38, Taylor and Francis, London.

Manacorda, M. (2012), "The Cost of Grade Retention", *Review of Economics and Statistics*, Vol. 94, No. 2, pp. 596-606, MIT Press, Cambridge, MA.

McCombs, M. and **D. Shaw** (1972), "The Agenda-Setting Function of Mass Media" *Public Opinion Quarterly*, pp. 176-187, Oxford University Press, Oxford.

McEwan, P. (2014), "Improving Learning in Primary Schools of Developing Countries: A Meta-analysis of Randomized Experiments", *Review of Educational Research*, Vol. 85, No. 3, pp. 353-394, Sage, New York, http://dx.doi.org/10.3102/0034654314553127.

Murnane, R. and **A. Ganimian** (2014), "Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations", *NBER Working Paper,* No. 20284, National Bureau of Economic Research, Cambridge, MA.

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

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

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

OECD (2011), *Lessons from PISA for the United States, Strong Performers and Successful Reformers in Education*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264096660-en.

OECD (2010), *Lessons from PISA for Mexico, Strong Performers and Successful Reformers in Education*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264107243-en.

OECD (2008) *The Global Evaluation of the Policy Impact of PISA*, Paper presented at the PISA Governing Board Meeting, The Hague.

Olper, A. and **J. Swinnen** (2013), "Mass Media and Public Policy: Global Evidence from Agricultural Policies", *World Bank Economic Review*, Vol. 27, No. 3, pp. 413-436, Oxford University Press, Oxford.

Parks, B., Z. Rice and **S. Custer** (2015), *Marketplace of Ideas for Policy Change: Who do Developing World Leaders Listen to and Why?,* AidData and The College of William and Mary, Williamsburg, VA.

Rivas, A. (2015), *Latin America after PISA: Lessons Learned about Education in Seven Countries,* Buenos Aires, CIPPEC-Natura-Instituto, Natura.

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

Schwerdt, G. and **M. West** (2012), "The Effects of Early Grade Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida", *Program on Education Policy and Governance Working Papers Series*, PEPG 12-09, Program on Education Policy and Governance, Harvard University.

Shadrova, A. (2015), "Frequency-of-mention of Selected Education-related Terms in Interaction with PISA Study Result Release Dates: Results from Brazil, Bulgaria, Colombia, Georgia, Indonesia, Jordan, Kyrgyzstan, Turkey and Vietnam", background paper prepared for "The Experience of Middle-Income Countries Participating in PISA 2000-2015".

Shamatov, D. (2014), "Education Quality in Kyrgyzstan and the Programme for International Student Assessment (PISA)", in D.B. Napier (ed.) *Qualities of Education in a Globalised World*, Springer.

Shamatov, D. and **K. Sainazarov** (2010), "The Impact of Standardized Testing on Education Quality in Kyrgyzstan: The Case of the Program for International Student Assessment (PISA) 2006", in A.W. Wiseman (ed.), *The Impact of International Achievement Studies on National Education Policymaking*, Vol. 13, Emerald Group Publishing Limited.

Simmons, J. and **L. Alexander** (1978), "The Determinants of School Achievement in Developing Countries: A Review of the Research", *Economic Development and Cultural Change*, pp. 341-357, University of Chicago Press, Chicago.

Strömberg, D. (2001), "Mass Media and Public Policy", *European Economic Review*, Vol. 45, pp. 652-663, Elsevier, Amsterdam.

Wagemaker, H. (2014), "International large-scale assessments: From research to policy", in L. Rutkowski, M. von Davier. and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues andMmethods of Data Analysis*, Taylor and Francis, New York.

World Bank (2012), *How Can Bulgaria Improve Its Education System? An Analysis of PISA 2012 and Past Results,* World Bank, Washington, DC, https://openknowledge.worldbank.org/handle/10986/20776, License: CC BY 3.0 IGO.

World Bank (2011), *Education Sector Strategy 2020: Learning for All*, World Bank Human Development Network, Washington, DC.

World Bank (1999), *Education Sector Strategy*, World Bank Human Development Network, Washington, DC.

World Bank (1995), *Priorities and Strategies for Education: A World Bank Review*, World Bank, Washington, DC.

World Bank (1990), *Primary Education: A World Bank Policy Paper*, World Bank, Washington, DC.

World Bank (1980), *Education Sector Policy Paper*, World Bank, Washington, DC.

PISA FOR DEVELOPMENT

6

What does PISA data tell us about education in middle-income countries?

This chapter reviews the evidence regarding the quality and equity of education in middle-income countries, as shown through the Programme for International Student Assessment (PISA) 2012 assessment of mathematics. The first section documents the extent of educational quality and inequality in the mathematics performance of students in 18 middle-income countries that participated in PISA 2012, as presented in OECD publications. The second section reviews the empirical evidence for systemic institutional factors (accountability, autonomy competition, tracking and preschool) that are related to higher levels of performance, generally, and as they are related to performance in the 18 countries participating in PISA 2012. It also examines school factors (school inputs, teacher quality, and instructional time) generally related to student performance in low- and middle-income countries and in the 18 middle-income countries participating in PISA 2012. The third section reports the results from new multi-level analyses and other statistical approaches undertaken for this chapter that further explore these topics.

INTRODUCTION

Successful education systems simultaneously achieve high levels of student achievement and an equitable distribution of achievement across all students, without regard to the immutable characteristics of their school (such as its geographical location) or of the student themselves (such as their social background, gender, ethnicity or home language). How educational systems become successful has been greatly debated over the past half century, and cross-country studies have been undertaken to shed light on this issue. As Foshay noted in 1962: "If custom and law define what is educationally allowable within a nation, the educational systems beyond one's national boundaries suggest what is educationally possible" (Hanushek and Wößmann, 2014).

"What is possible" includes systemic possibilities as well as school, classroom and teacher possibilities. These have all been explored extensively across the typically high-income countries/economies that have participated in international large-scale assessments in the 1970s, 1980s, 1990s and early 2000s. They have also been explored to a lesser extent in regional large-scale assessments. Countries that have performed less well on these assessments have often turned to higher performing countries/economies for advice. But the growing number of middle-income countries participating in international large-scale assessments raises the question of how much the experience of higher-income countries can be generalised across these new boundaries. The purpose of this chapter is to explore this issue.

The most recent PISA reports from PISA 2012 provide a massive amount of information – over 2 000 pages of text and data – regarding the education systems that participated in the assessment. This chapter draws, in part, from four volumes:

1. *PISA 2012 Results Volume I: What Students Know and Can Do: Student Performance in Mathematics, Reading and Science* (OECD, 2014).
2. *PISA 2012 Results Volume II: Excellence through Equity: Giving Every Student the Chance to Succeed* (OECD, 2013a).
3. *PISA 2012 Results Volume III: Ready to Learn: Students' Engagement, Drive and Self-Beliefs* (OECD, 2013b).
4. *PISA 2012 Results Volume IV: What Makes Schools Successful? Resources, Policies and Practices* (OECD, 2013c).

It is beyond the scope of this chapter to summarise all the analyses in these volumes, but in the sections below some key findings relevant to middle-income countries are reported. In particular, issues of equality in opportunity linked to students' socio-economic background are explored at length.

This chapter is organised as follows. First, it documents the extent of educational quality and inequality in middle-income countries, with particular reference to the mathematics performance of students in 18 middle-income countries that participated in PISA 2012.[1] Then, it reviews the empirical evidence for systemic and school (including classroom and teacher) factors that are associated with higher levels of performance in general, and with lower performance gaps between groups. This draws on recent literature and the OECD's analysis of PISA 2012, with specific reference made to the 18 participating middle-income countries. Finally, it presents the results from original analyses of the data from these 18 countries that were carried out for this report.[2]

EDUCATIONAL QUALITY AND INEQUALITY IN MIDDLE-INCOME COUNTRIES

Quality is generally lower than in OECD countries

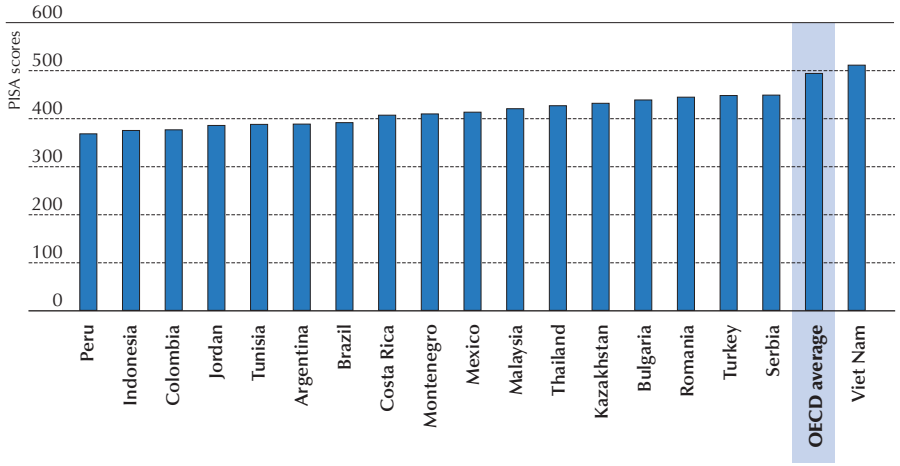
The overall quality of learning outcomes in low- and middle-income countries has been studied extensively in recent decades, drawing on various international large-scale assessments, regional large-scale assessments, and national assessments. The general finding is that the quality of learning outcomes in low- and middle-income countries, assessed at every level from primary school through to upper secondary school, is often very poor. However, variations in learning outcomes are large, and there exist well-performing low- and middle-income countries and well-performing schools within these countries.

The OECD analyses reported in the PISA 2012 volumes show that the overall performance of 15-year-old students in all of the middle-income countries participating in PISA 2012, other than Viet Nam, was lower than that of students in the OECD countries, and varies widely (Figure 6.1). Average mathematics scores in eight countries – Albania, Argentina, Brazil, Colombia, Indonesia, Jordan, Peru and Tunisia – fell below 400 points – compared with the OECD average of 494 points (OECD, 2013c).[3]

The highest-performing country from this group was Viet Nam (511 points), which exceeded the OECD average by 17 points. Other relatively well-performing countries in the pool of below-average performers were: Bulgaria (439), Romania (445), Turkey (448), and Serbia (449). In most middle-income countries, performance in PISA has been steadily increasing over the past decade. The average annual improvement has been largest in Albania, Kazakhstan and Malaysia, at more than 5 points per year. Brazil, Bulgaria and Romania increased their annual performance by more than 4 points per year.

Figure 6.1

Average scores in mathematics in 18 low- and middle-income countries, PISA 2012

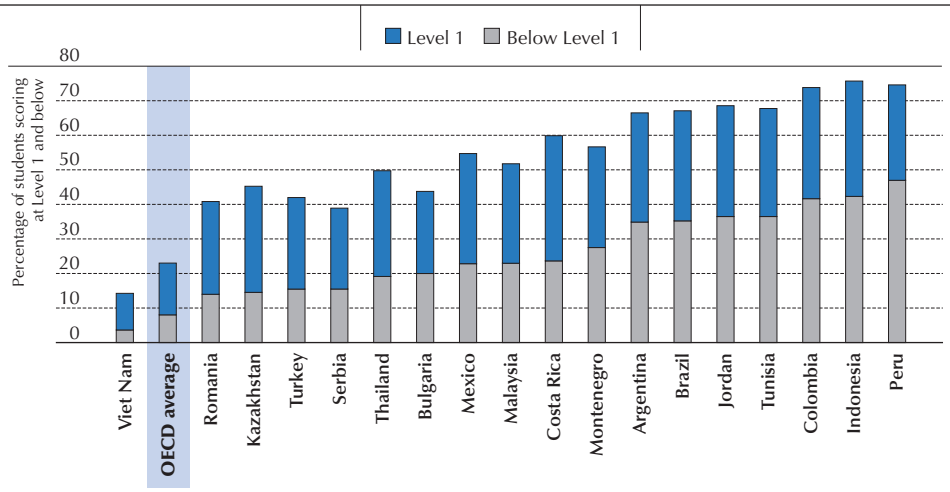


Source: OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en.

StatLink http://dx.doi.org/10.1787/888933293906

Figure 6.2

Percentage of students scoring at level 1 and below in mathematics, PISA 2012



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

StatLink http://dx.doi.org/10.1787/888933293910

PISA scores have been divided into six proficiency levels, from low proficiency (Level 1) to high proficiency (Level 6); a definition of each level is provided in Annex D. Low scores represent a concentration of students with low mathematics proficiency. Across OECD countries: 23% of students score at or below Level 1; 22.5% score at Level 2 (from 420 to less than 482); 23.7% score at Level 3 (from 482 to less than 544); 18.1% score at Level 4 (from 544 to less than 606); 9.3% score at Level 5 (from 606 to less than 669); and 3.3% score at Level 6 (above 669 points) (OECD, 2014).

For more than half the participating middle-income countries, scores on the PISA mathematics test are concentrated at Level 1 or below (Figure 6.2). Approximately 75% of students in Colombia, Indonesia and Peru, and over 60% of students in Albania, Argentina, Brazil, Jordan and Tunisia, score at or below Level 1 in mathematics. In all of these countries only 1% of students are top performers, compared with 12.6% of students in OECD countries. This distribution of PISA scores in middle-income countries is addressed in a recent paper on the enhancement of PISA cognitive instruments (Adams and Cresswell, 2014).

One consequence of low PISA scores in many middle-income countries is a low variation in scores. Half of the middle-income countries in this study have the narrowest spread in student mathematics scores among all PISA participants; these countries are: Argentina, Brazil, Colombia, Costa Rica, Indonesia, Jordan, Kazakhstan, Tunisia and Thailand. Among the remaining nine countries where performance was higher, the variation in mathematics scores in Serbia and Turkey was close to the OECD average, while the variance in scores in Bulgaria exceeded the OECD average. Viet Nam is the only participating middle-income country that combines above-average PISA performance with below-average variation in performance.

Inequality is generally greater in middle-income countries than in OECD countries

Inequalities in educational outcomes are found in all countries and at all income levels. The sources of these inequalities differ from country to country, but they are often due to: the socio-economic background of the family (including parental education and occupational status), gender, ethnicity (including a disjoint between the language spoken at home and the language of school instruction) and school location (urban versus rural). These student and school background characteristics are frequently referred to as a "disadvantage", and in many countries some children, often girls, face a "double disadvantage" when two or more of these disadvantages are combined (Lewis and Lockheed, 2007).

The specific characteristics of inequalities are likely to vary across countries. Table 6.1 describes some differences among the middle-income countries that participated in PISA 2012, with respect to several student and school characteristics.

Table 6.1
Student characteristics, PISA 2012

	In rural schools (%)	Female (%)	Non-speaker of language of assessment (%)	In grade 7 or grade 8 (%)	Has repeated a grade (%)	Socio-economic status
OECD Average	9.4	--	10.6	5.4	12.4	0
Argentina	8.5	51.4	1.6	14	36.2	-0.72
Brazil	1.1	52.2	1.1	6.9	36.1	-1.17
Bulgaria	3.8	48.2	10.8	5.5	4.8	-0.28
Colombia	13.0	52.9	0.7	17.6	40.6	-1.26
Costa Rica	23.5	53.1	1.2	21.1	33.5	-0.98
Indonesia	29.1	49.2	58.9	10.2	15.5	-1.8
Jordan	10.4	50.6	4.7	1.2	7.9	-0.42
Kazakhstan	32.9	50.2	11.1	5.1	1.6	-0.32
Malaysia	13.4	51.6	42.3	0.1	0	-0.72
Mexico	15.1	51	3.2	6.3	15.5	-1.11
Montenegro	0	50	1	0.1	1.3	-0.25
Peru	18.8	51.4	6.4	10.5	27.5	-1.23
Romania	8.2	51	1.2	7.6	4.5	-0.47
Serbia	0.4	50.2	4.2	1.6	1.6	-0.3
Thailand	15.7	56	44.6	0.4	3.3	-1.35
Tunisia	4.4	53.4	1.1	16.8	38.7	-1.19
Turkey	2.3	49.5	6.2	2.7	14.4	-1.46
Viet Nam	45	--	2.2	3.1	7.7	-1.81

Sources: OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en.
OECD (2013a), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201132-en.
OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201156-en.
StatLink http://dx.doi.org/10.1787/888933294069

A higher share of students in middle-income countries attend schools in rural locations, defined by the OECD as "a village, hamlet or rural area with fewer than 3 000 people", compared with students in OECD countries. In OECD countries, about 9% of students live in rural communities; this share is much larger in several middle-income countries. A slightly higher share of students in middle-income countries are female compared to OECD countries, but across all countries girls account for about half the students in the PISA assessment. Students who speak a home language that differs from the language of instruction – and, by inference, the language of the assessment, which is the measure of language differences used by PISA – are also more often found in middle-income countries than in OECD countries. In particular, the OECD average for students who do not speak the language of assessment at home is about 10%, whereas the share in some middle-income countries is several times higher: 59% in Indonesia, 45% in Thailand, and 42% in Malaysia.

In six countries the share of 15-year-olds in grades 7 or 8 is twice as high or greater than the OCED average. In Costa Rica, approximately 20% of 15-year-olds are in these grades. In part, this is due to the much higher rates of repetition, particularly in Latin American countries and Tunisia, where on average about one-third of students have repeated at least one year of schooling. It may also be due to a later school starting age in some countries.

The OECD uses an index of economic, social and cultural status, which comprises indicators such as parental education, parental occupation, cultural possessions in the home, educational resources in the home, and information and communication technology in the home. The index is calculated so that the OECD average is zero and the standard deviation is 1.0. The index values for all middle-income countries is below zero, ranging from 1.8 standard deviations below the OECD average in Viet Nam and Indonesia to 0.25 standard deviations below the OECD average in Montenegro. The value of this index varies widely within middle-income countries, with the socio-economic background of students in the lowest quartile of the index falling three standard deviations below the country average in some countries, and rising to one standard deviation above for students in the highest quartile of the socio-economic index in one country. In short, the students assessed by PISA in the middle-income countries differ substantially from those assessed in the OECD countries, and on average are more disadvantaged.

OECD ANALYSES OF CORRELATES OF PERFORMANCE

Multi-country international large-scale assessments can reveal differences among countries that are associated with variations in learning outcomes and have advantages over single-country studies (Hanushek and Wößmann, 2014). Since international large-scale assessments involve education systems with different institutional and school features, they provide a useful setting for understanding the link between variations in these features and educational outcomes. The OECD has conducted thorough analyses of the PISA 2012 data, only some of which is reported in this chapter. The OECD has examined: *i)* the quality and equity in the outcomes of PISA 2012, by country, in Volume II of the results from PISA 2012; *ii)* student attitudes and motivations in Volume III; and *iii)* various school-level correlates of student performance, by country, in Volume IV. This section reviews the correlates of performance at three levels: between-country correlates, between-school correlates within countries, and between-student correlates within schools and countries.

Systemic differences among countries

Much of the literature on the effects of systemic differences on student learning has focused on systemic differences among countries with respect to institutions and educational inputs.

Institutions and learning in OECD countries

A recent review article by Hanushek and Wößmann (2014) summarises the findings from 19 cross-country studies from the late 1990s and early 2000s, in which the majority of the countries were high-income economies. They identify five institutional features that differentiate the education systems among these high-income countries: accountability measures, school autonomy, competition and private involvement, early school tracking and the pre-primary education system. These features often create variations in educational incentive systems.

School accountability in the form of curriculum-based exit exams and regular standardised testing improves learning by increasing the reward of learning for students while incentivising stakeholders to monitor learning (Wößmann, 2003, 2005).

School autonomy is found to have positive effects on learning, but the positive impact from decentralised level decision makers, who have better access to information on best practices, only occurs if there is alignment between the incentives for schools and those for students (Fuchs and Wößmann, 2007).

School competition, mostly in the form of competition between private and public schools, can improve the performance of both public and private schools under certain conditions (Wößmann, 2007). The advantage of private schools is higher in the countries where private schools receive large shares of public funding. This is assumed to be the consequence of a level playing field between private and public schools, through which competition and self-selection into these schools changes the mentality of the students, parents and the school staff.

Tracking is a way of placing students in schools and/or classes according to their past or anticipated performance (both academic and non-academic) and has been found to increase inequality (Wößmann, 2009). More homogenous classes may or may not help create optimal learning situations depending on the nature of the peer dynamics. While the optimisation of learning will take place through tailor-made curricula, tracking will, if implemented early on, disadvantage the weaker groups of students and increase the inequality of student achievement.

Pre-primary education enhances early learning and can potentially mitigate the effect of family background for disadvantaged children. Structural quality and the access to pre-primary education are positively related to both performance and greater equality of the educational systems (Berlinski, Galiani and Gertler, 2009; Engle et al., 2007; Schutz, Ursprung and Wößmann, 2008).

Inputs and learning in OECD countries

Educational systems differ in the amount and type of resources allocated to students at different levels, including resources such as per-student expenditure, quality of teachers and the amount of time for learning. Early research – with data primarily from high-income countries/economies – concluded that cross-national differences in educational expenditure and instructional time were generally unrelated to cross-national differences in performance in mathematics (Baker, Goesling and LeTendre, 2002; Hanushek, 1997). By comparison, one recent 46-country cross-national study on the effects of teacher quality, as measured by certification or degree, found a positive relationship to mathematics achievement across these countries (Akiba, LeTendre and Scribner, 2007). An analysis of PISA 2003 results in mathematics in 29 OECD countries concluded that expenditure per student, instructional time and teacher quality were all related to differences in achievement, but that the effects were small (Hanushek and Wößmann, 2014).

Differences between effects in OECD and middle-income countries

Can lessons derived from these analyses of systemic differences across high-income countries translate directly into advice for low- and middle-income countries? The answer is most likely "not automatically". Most cross-country research does not disaggregate results according to the country's economic status nor take into account associated differences in the cultural and socio-economic context between high-income and lower-income countries. Studies that examine growth in average student achievement, in particular, often fail to recognise that educational policy in low- and middle-income countries has focused on improving access (improving enrolment, attendance and school completion rates), stimulated by the education-related Millennium Development Goals (UNESCO, 2015). This has resulted in greater inclusiveness in the composition of the student population, but can also lower a country's average performance, due to the higher share of disadvantaged students in the overall student population.

Some system-level institutions associated with improved performance in high-income countries/economies are associated with lower levels of performance in middle-income countries/economies. Using the PISA 2003 data, Hanushek, Link and Wößmann (2013) show that the impact of school autonomy on student achievement is highly heterogeneous and varies across countries and educational systems. At low levels of economic development, increased school autonomy – particularly in the decision-making areas related to instructional content, but also in the areas of personnel and budgeting – is associated with lower student outcomes. One reason for this may be the lack of interaction and co-operation between the overall institutional structures in developing countries and schools and their leadership. If the overall institutional structure is weak, the danger of decentralisation lies in the possibility that individual schools, often with weak leadership and little feedback regarding the consequences of decisions, pursue poorly chosen goals.

Similarly, the OECD (OECD, 2013c) shows that educational spending is not related to average performance at the country level across developed countries, yet this is definitely not the case for low- and middle-income countries, where the level of expenditure on education per students is very strongly correlated with educational performance. Hence, while money alone does not buy strong performance for developed countries, it certainly has very high returns in developing countries, where schooling systems are embedded in very different socio-economic realities. How the incentive structures guide the use of resources is seen as the most important determinant of educational outcomes for developed countries, but the lack of basic resources and the inequality of access to these resources is a key problem in low- and middle-income countries.

The OECD's analyses of the five features of education institutions (detailed above) identified by Hanushek and Wößmann (2014) provide support for the claim that accountability, autonomy, competition and tracking have little effect on student performance in middle-income countries, once the socio-economic background of students and the socio-economic composition of schools have been taken into account (Table 6.2). However, these institutional features are measured within countries and are based on school-level reports.

The OECD's within-country multi-level analyses show few positive associations between these institutional features and student performance in PISA 2012 mathematics in the participating middle-income countries, as summarised below and in Table 6.2:

- **Accountability systems** (external evaluation, student feedback and public posting of achievement results) are each related to student performance in two countries, with both positive and negative effects for external evaluation and public posting of results.
- **School autonomy** for resources is unrelated to performance in 16 countries and negatively related to performance in 2 countries, while autonomy for curriculum and assessment is positively related to performance in 2 countries.
- **Competition** is unrelated to student performance in all countries, but private schools have higher achievement in two countries and lower performance in two countries.
- **Ability grouping** is negatively related to performance in two countries whereas academic selection is positively related to performance in one country.
- **Preschool** (the percentage of students in a school who had attended) is positively associated with performance in six countries and negatively in one.

— Table 6.2 —

Institutional features and mathematics achievement in 18 middle-income countries, PISA 2012

Institutional features	PISA measure	Countries where feature has significant positive effect on mathematics achievement	Countries where feature has significant negative effect on mathematics achievement
Accountability	External evaluation: Administrative authority tracks data over time	Thailand	No countries
	Feedback from students	Colombia, Viet Nam	No countries
	Achievement data are posted publicly	Bulgaria	Mexico
Autonomy	Over resources	No countries	Bulgaria, Indonesia
	Over curriculum and assessment	Costa Rica, Thailand	
Competition	With other schools	No countries	No countries
	With private schools	Mexico, Viet Nam	Colombia, Indonesia
Tracking	Ability group in mathematics	No countries	Bulgaria, Turkey
	Academic selection	Turkey	No countries
Preschool	Percentage of students who attended pre-primary	Bulgaria, Malaysia, Mexico, Romania, Thailand, Viet Nam	Jordan

Note: Results are based on OECD's multilevel analysis and analyses control for socio-economic composition of the schools and socio-economic background of the students.

Source: OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201156-en (Table IV.1.12c).

Differences among schools within countries

Unlike studies of systemic features, which have rarely looked specifically at low- or middle-income countries/economies, a great deal of literature on school features and their correlates with student performance within low- and middle-income countries has developed over the past several decades. This literature has been summarised in a number of recent reviews that address the quality of the research and the lessons learned (Glewwe et al., 2014; McEwan, 2014; Murnane and Ganimian, 2014). The conclusions from these reviews differ somewhat from conclusions reached in meta-analyses of "what works" in developed countries (Hattie, 2013; Wang, Haertel and Walberg, 1990), but are relatively consistent with the conclusions reached by Lockheed and Verspoor in an earlier review on school and family inputs (1990). Much of this research focuses on learning at the primary level, since until recently a very high share of secondary level age students had either still been in primary school or had been out-of-school altogether.

School inputs matter in many countries

In the late 1980s, Lockheed and Verspoor (1990) examined school inputs (curriculum, instructional materials, learning time and teaching quality) and family inputs (preschool experience and health and nutrition) as direct determinants of student learning in developing countries. Because randomised-control trials (RCTs) had been carried out for very few school inputs in few countries, most of the studies that were reviewed used cross-sectional correlational analyses, or "education production functions" (Fuller, 1987; Heyneman and Loxley, 1983; Lockheed and Hanushek, 1988; Simmons and Alexander, 1978). At the time it was widely thought that school inputs were more important than home background in developing countries, a departure from the evidence in developed countries. With the advent of new algorithms for analysing multi-level, hierarchically organised data, these conclusions about the relative importance of school and family background were called into question (Raudenbush and Willms, 1991; Riddell, 1989). During the 1990s, a range of new studies were carried out, often using multi-level analyses of cross-sectional data (Riddell, 2008).

Contemporary research identifies school inputs as important, but also acknowledges both the role of families and the broader institutional factors. Many recent reviews focus on studies that have used an experimental design (RCTs) to test their hypotheses, a sharp departure from the correlational analyses used in earlier research. Although RCTs are widely regarded as a "gold standard" for research methods, their use in evaluating school inputs has been limited to very few school inputs in low- or middle-income countries, and the conclusions from the few RCTs that have been carried out are inconsistent (Evans and Popova, 2015). Correlational analyses offer the opportunity to examine many additional school features.

Recent reviews of research on school inputs cover research using both correlational and experimental designs. In a review of 79 "good quality" studies, including 43 identified as "high quality" with respect to methodology, Glewwe et al. (2014) consider how physical school infrastructure and pedagogical supplies, characteristics of the principal and financial resources, contribute to student learning in developing countries, after the effects of student background has been statistically controlled. McEwan's (2014) review of experiments examines school inputs (instructional inputs, teacher capacity, teacher effort, instructional time and health inputs) as well as parental inputs. Willms and Tramonte's (2015) review adds to these inputs a child's early learning opportunities, home and school language. Murnane and Ganimian (2014) draw greater attention to the teacher and teaching process. Evans and Popova (2015) observe that the overlap in studies examined in six reviews, including those of McEwan, Willms and Tramonte, and Murnane and Ganimian, is "surprisingly limited" with 75% of the studies occurring in only one of the reviews.

Despite the differences among these reviews and meta-analyses in terms of the types of research and the topics covered, some of the conclusions are remarkably similar:

- **Basic school inputs** – textbooks, desks and chairs, electricity, and even the quality of the school's roof and walls – make a difference in many low- and lower-middle-income countries where such inputs are not widely available (Glewwe et al., 2014).
- **Teacher quality,** particularly subject matter knowledge and teaching practice, is consistently related to student learning in low- and middle-income countries (Glewwe et al., 2014; Kremer, Brannen and Glennerster, 2013; McEwan, 2014; Murnane and Ganimian, 2014) as well as in high-income countries (Hattie, 2013).
- **Instructional time** – greater when teacher absenteeism is lower – is also related to student learning in some countries (Glewwe et al., 2014; Lavy, 2010; Long, 2013).

Where measured, however, the effect sizes for most school-level inputs are quite low, typically less than 15% of a standard deviation of the learning outcome measure. However, the effect sizes for teacher quality and teaching practice are quite high, amounting to over 40% of a standard deviation of learning outcomes measured in high-income countries (Hattie, 2013).

The OECD PISA instruments measure only a few of these items. Educational resources and quality of infrastructure are reported in two indices derived from the school questionnaire regarding shortages that are perceived to affect student performance; for analyses, the indices are reversed so that a higher number represents a more positive learning environment. Educational resources include science laboratory equipment, instructional materials, computers, Internet, computer software and library materials. Physical infrastructure includes buildings and grounds, heating/cooling and lighting, and classrooms. Teacher quality is measured in three ways: the proportion of teachers that are certified, the proportion of teachers who have higher education (ISCED 5A) and the proportion of teachers who have received recent

in-service training in mathematics. Instructional time is measured through student reports regarding the average length in minutes of class periods, and the number of class periods per week for mathematics, languages and science; these are combined as a measure of formal learning time for each subject.

Shortages in these key inputs are prevalent in various countries (Table 6.3). For example, in Colombia, Costa Rica, Peru and Tunisia, more than half of students attend schools where insufficiencies in instructional materials are reported. In several countries, the weekly time for regular mathematics lessons falls well below the OECD average, and students in middle-income countries report having covered less formal mathematics than students in OECD countries.

Table 6.3

Selected educational resource inputs in 18 middle-income countries, PISA 2012

	Teacher quality (% with mathematics in-service training)	Teacher quality (% with ISCED 5A)	Instructional materials (% no-shortage)	Time (weekly minutes for mathematics lessons)
OECD Average	39	88	80	218
Argentina	48	18	62	267
Brazil	36	87	86	215
Bulgaria	36	--	75	134
Colombia	22	91	33	263
Costa Rica	46	84	43	207
Indonesia	42	82	62	209
Jordan	33	85	74	227
Kazakhstan	36	85	53	182
Malaysia	43	89	93	201
Mexico	47	88	60	253
Montenegro	46	89	60	142
Peru	33	77	42	287
Romania	45	96	71	169
Serbia	48	7	51	154
Thailand	73	99	63	206
Tunisia	40	87	41	276
Turkey	18	93	72	172
Viet Nam	50	87	73	227

Source: OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing, Paris, http://dx.doi.org/ 10.1787/9789264201156-en (Figure IV.3.8, Tables IV.3.21, IV.3.6 and IV.3.12).

StatLink http://dx.doi.org/10.1787/888933294072

The results from the OECD multi-level analyses, which control for both student-level and school-level socio-economic status, are summarised in Table 6.4 for 18 middle-income countries. In general, the inputs as measured by the OECD instruments are unrelated to differences in performance across schools in the majority of these middle-income countries, despite reported inadequacies of instructional materials such as textbooks and libraries in several countries. The physical infrastructure of the school is unrelated to performance in 15 countries and negatively related to performance in 3 countries.

Table 6.4

School inputs and mathematics performance in 18 middle-income countries, PISA 2012

Education Inputs	PISA measure	Countries with positive correlation with performance	Countries with negative correlation with performance
Physical infrastructure	Index of quality of physical infrastructure	No country	Jordan, Romania, Turkey
Educational resources	Index of quality of educational resources	Costa Rica, Romania	No country
Teacher quality	Proportion of teachers with ISCED 5A	No country	Peru, Romania
	Proportion of teachers having attended professional development	Argentina, Malaysia	No country
Instructional time	School average of students' learning time per week	Argentina, Kazakhstan, Malaysia, Mexico, Turkey	Brazil

Note: Results from multi-level analyses, controlling for student and school demographics and socio-economic status; only statistically significant results are reported.

Source: OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing, Paris, http://dx.doi.org/ 10.1787/9789264201156-en (Table IV.1.12c.).

Educational resources are unrelated to performance in 16 countries and positively related to performance in Costa Rica and Romania. Teacher quality (university degree) is unrelated to performance in 16 countries and negatively related to performance in 2 countries, while in-service training is positively related to performance in 2 countries. Instructional time is unrelated to performance in 12 countries, positively related to performance in 5 countries, and negatively related to performance in 1 country.

Opportunity to learn and school climate matter in most countries

Studies in OECD countries regarding "what works" to raise student learning emphasise the importance of the opportunity to learn what is being taught (Schmidt, Zoido, and Cogan, 2014). To measure exposure to different contents of learning – which is typically referred to as the "implemented curriculum" – the PISA 2012 assessment asked students to recall their exposure to mathematical theories, concepts and content, and the amount of class time they spent studying this content.[4] The answers to these questions were used to create three indices that measure opportunities to learn at the student level: "formal mathematics", "word problems" and "applied mathematics." Of these three sub-indicators of opportunities to learn, exposure to and experience with formal mathematics was found to be the strongest correlate of mathematics performance (Schmidt et al., 2014). The level of exposure to formal mathematics is relatively high in some of the low- and middle-income countries (Bulgaria, Jordan, Montenegro, Romania, Serbia and Viet Nam) even by OECD standards. In others, it is low (Argentina, Brazil and Tunisia).

PISA 2012 assessed three aspects of the school climate: *i)* degree of discipline among students; *ii)* the quality of relationships between students and their teachers; and *iii)* the values promoted and shared between teachers and students, and among the students themselves. Disciplinary climate was highly correlated with performance across all countries, with few exceptions (OECD, 2013c). To measure the disciplinary climate, PISA 2012 asked students to identify the frequency with which interruptions occur in mathematics lessons by indicating how often: students do not listen to what teachers have to say, there is noise and disorder, teachers have to wait a long time for students to calm down, students cannot work well, and students do not start working for a long time after the lesson begins. These responses were combined in the composite "index of disciplinary climate," with a mean of zero and a standard deviation of 1 for OECD countries. Higher values indicate that students perceived a positive disciplinary climate in their classrooms, whereas a negative value indicates that students perceived a negative disciplinary climate in their classrooms. Among the middle-income countries, the indices were positive in 7 (Costa Rica, Indonesia, Kazakhstan, Mexico, Romania, Thailand, Viet Nam), and negative in 11 (Argentina, Brazil, Bulgaria, Colombia, Jordan, Malaysia, Montenegro, Peru, Serbia, Tunisia, Turkey) (OECD, 2013c).

Teaching practices and teacher's knowledge matter

The meta-analysis literature on "what works" in high-income countries has consistently identified specific teaching practices as having the highest associations with student learning (Hattie 2013; Walberg 1984). Such influences include: teachers providing feedback to students, the quality of instruction, direct instruction, assigning homework and questioning; all these influences have effect sizes over 0.40. Where information has been available, prior research has documented the positive impact of teacher knowledge on student learning outcomes (Murnane and Ganimian, 2014). In the past, PISA instruments have neither included a teacher questionnaire to investigate teaching practices nor measured teacher subject matter knowledge. Recently, PISA 2012 results have been linked with the OECD's Teaching and Learning International Study (TALIS) for six high-income countries, Mexico and Romania (Austin et al., 2015). For PISA 2015, a teacher questionnaire will be included.

Differences among students between and within schools

The OECD has recognised the importance of a student's home background, and all OECD multi-level analyses have adjusted for these effects. Specifically, most analyses have introduced statistical controls for a student's gender, home language and socio-economic background. In addition, the analyses control for a school's urban or rural location, which can be considered a reasonably proxy for a student's residence as well.

These student demographic characteristics are strongly correlated with performance in most of the middle-income countries analysed by the OECD (Table 6.5). For all countries in PISA 2012, a student's socio-economic background was strongly and significantly associated with reading, mathematics and science performance. For simplicity, only the results for mathematics performance are presented in the table.

The OECD reports substantial performance gaps associated with many of the fixed characteristics of schools and students (Table 6.5). Students in urban schools outperformed students in rural schools in virtually all middle-income countries, and the gap exceeded the OECD average in about half of the countries. In most cases, these gaps amount to one or two

years of schooling. Only in Turkey did the scores of students in rural areas exceed those in urban areas. This could be due to two types of selection effects: *i)* the very small share of students attending school in rural areas in Turkey (2.3%, which is well below the OECD average); and *ii)* the policy of locating academically elite "Anatolia" high schools in all regions of the country, including rural areas.

Table 6.5

Student background and mathematics performance, PISA 2012

Socio-economic and demographic student background	PISA measure	Significantly higher performance in:	Significantly lower performance in:
Gender	Female	Jordan, Malaysia, Thailand	Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Kazakhstan, Mexico, Peru, Serbia, Tunisia, Viet Nam
Home language	Student speaks language of assessment at home (non-immigrants)	Argentina, Bulgaria, Mexico, Peru, Romania	Indonesia, Malaysia
Socio-economic background	OECD index of family socio-economic status	Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Indonesia, Jordan , Kazakhstan, Malaysia, Mexico, Montenegro, Peru, Romania, Serbia, Thailand, Tunisia, Turkey, Viet Nam	No country

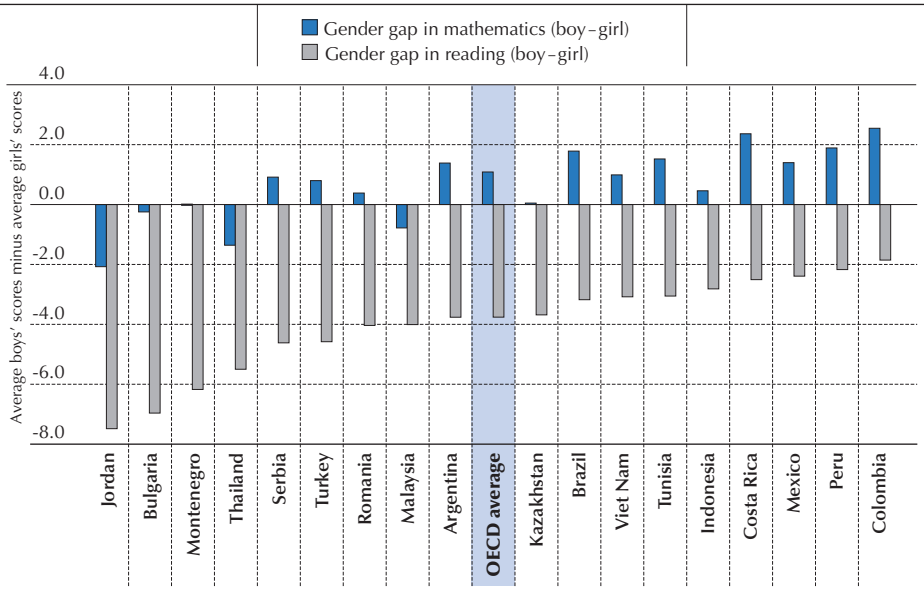
Sources: OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en (Table I.2.3a [gender]).
OECD (2013a), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201132-en (Table II.3.5 [language spoken at home], Table II.2.1 [socio-economic status]).

StatLink http://dx.doi.org/10.1787/888933294084

Gender effects differ for reading and mathematics. In many countries, boys outperformed girls in mathematics, whereas girls outperformed boys in reading (Figure 6.3). Girls significantly outperformed boys in reading performance in the middle-income countries In more than half of these countries the female reading advantage even exceeded the female reading advantage in OECD countries. By comparison, boys outperformed girls in mathematics in OECD countries, although the statistically significant difference is small (11 points). In all the Latin American countries participating in PISA 2012, however, this difference in favour of boys in mathematics was quite substantial (Figure 6.3). In Jordan, Thailand and Malaysia, girls outperformed boys in mathematics.

Figure 6.3

Gender gap in mathematics and reading in 18 middle-income countries, PISA 2012



Source: OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264208780-en (Table 1.2.a).

StatLink http://dx.doi.org/10.1787/888933293925

Home language that differs from the language of instruction was associated with lower mathematics scores in about half of participating countries and with higher mathematics scores in Indonesia and Malaysia. In several Latin American countries, these performance differences may reflect large indigenous ethnic minorities. In Bulgaria and Romania – two other countries with large language gaps – a large share of Roma may account for the gaps. In Malaysia and Indonesia, students who did not speak the language of assessment at home actually outperformed those who did. The OECD discusses language minority groups in the context of immigrants. In low- and middle-income countries, however, language minorities are often indigenous peoples who may not speak the official national language in their homes. In OECD countries, fewer than 5% of students were non-immigrant students who reported speaking a language other than the language of assessment in their home. In the low- and middle-income countries, this share was as high as 59% in Indonesia, 42% in Malaysia, 44% in Thailand, and 15% in Kazakhstan.

Gaps between the performance of students with a higher socio-economic status and those with a lower socio-economic status occur both within schools and between schools and, in general, are smaller in middle-income countries than the average OECD country. The student-level score point difference associated with a one-unit increase in student-level socio-economic status, and the school-level score-point difference associated with a one-unit increase in the school mean socio-economic profile, are lower than those in the average OECD country, with a few differences. In Kazakhstan, Malaysia and Romania, the within-school gap approaches that of the OECD average, and in Montenegro, Serbia and Turkey, the between-school gaps are higher than the OECD average. The relatively low gaps may simply be an artefact of the distribution of scores in these countries.

Table 6.6

Gaps in mathematics performance by student characteristics in 18 middle-income countries, PISA 2012

	Gender (boy – girl)	Socio-economic status within school	Socio-economic status between schools	Language at home	Location (urban / rural)
OECD Average	11*	19*	72*	33*	36*
Argentina	14*	9*	49*	45*	31*
Brazil	18*	8*	46*	-2	41*
Bulgaria	-2	12*	73*	77*	107*
Colombia	25*	11*	35*	36*	50*
Costa Rica	24*	10*	34*	-19	36*
Indonesia	5	6*	37*	4	48*
Jordan	-21*	11*	47*	1	38*
Kazakhstan	0	15*	45*	-12	23*
Malaysia	-8*	15*	49*	-31*	63*
Mexico	14*	5*	29*	47*	58*
Montenegro	0	12*	102*	21	--
Peru	19*	10*	49*	76*	89*
Romania	4	17*	57*	31*	60*
Serbia	9*	9*	101*	8	--
Thailand	-14*	9*	35*	12*	35*
Tunisia	15*	6*	45*	11	40*
Turkey	8	6*	83*	52*	-39
Viet Nam	10	8*	49*	52*	62*

* $p < .01$.

Source: OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/ 9789264201156-en (Tables 1.2.3a, II.3.5, Figure II.5.1b, II.3.3.a, Table 1.3a, Table II.2.9a).

StatLink http://dx.doi.org/10.1787/888933294095

Disadvantaged students, disadvantaged schools

OECD analyses investigate how disadvantaged students often attend disadvantaged, under-resourced schools. These analyses are reported at length in *PISA 2012 Results, Volume IV*, and this section summarises three of the findings regarding school inputs: instructional time, physical infrastructure, and educational resources.

In the majority of middle-income countries, socially advantaged schools (defined as a school whose students' mean socio-economic status is statistically significantly above the country/economy mean) benefit from more instructional time, a better physical infrastructure and more educational resources compared with socially disadvantaged schools (a school whose students' mean socio-economic status is statistically significantly below the country/economy mean).

Table 6.7

Socio-economic school disadvantage and access to resources in 18 middle-income countries, PISA 2012

School resources	More resources in socially advantaged schools	No difference between socially advantaged and socially disadvantaged schools	More resources in socially disadvantaged schools
Learning time (per week in mathematics)	Argentina, Brazil, Bulgaria, Costa Rica, Indonesia, , Kazakhstan, Malaysia, Mexico, Montenegro, Peru, Serbia, Thailand, Tunisia, Turkey, Viet Nam	Jordan, Colombia, Romania	No country
Physical infrastructure	Argentina, Brazil, Colombia, Costa Rica, Indonesia, Mexico, Montenegro, Peru, Thailand, Turkey, Viet Nam	Jordan, Kazakhstan, Malaysia, Romania, Serbia, Tunisia	Bulgaria
Educational resources	Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Indonesia, Jordan , Malaysia, Mexico, Peru, Romania Thailand, Tunisia, Turkey, Viet Nam	Kazakhstan, Montenegro, Serbia	No country

Source: OECD (2013c), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/ 9789264201156-en (Table IV.3.22).

In none of the middle-income countries do schools with a higher share of disadvantaged students receive more resources than schools with a higher share of advantaged students. These differences appear to be related to early tracking into vocational programmes in some countries, and the prevalence of private education, particularly in Latin America and Thailand. A few countries – notably Kazakhstan and Serbia – appear to provide both infrastructure and instructional materials equitably, regardless of the socio-economic composition of students in the schools.

The OECD has carried out extensive multi-level analyses of factors related to student learning. With the exception of the student's gender, home language and socio-economic status, and the school's socio-economic status, very few of the explanatory variables in the OECD models are statistically significant in more than a few countries.[5] The results relating to school inputs and systemic differences have been summarised above. In addition, the OECD has examined a number of student attitudes and experiences, including the opportunity to learn material covered in the PISA assessment. The OECD's analysis, however, considers only a few of the variables in the PISA data set and tends to reflect indices constructed on the basis of questions developed for students and principals in OECD countries. Deeper analyses of PISA data may shed light on education in middle-income countries and policies that could lead to improvements in education in these countries. Such analyses are undertaken in the following section.

NEW ANALYSES: CORRELATES OF QUALITY AND SOURCES OF INEQUALITY

This section reports original analyses, undertaken specifically for this report, for the 18 middle-income countries that participated in the most recently completed PISA cycle, PISA 2012: Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Indonesia, Jordan, Kazakhstan, Mexico, Montenegro, Malaysia, Peru, Romania, Serbia, Thailand, Tunisia, Turkey and Viet Nam. Countries not included in this analysis are Albania and Hungary.[6] The analysis also does not include any of the economies in the Reople's Republic of China (hereafter 'China') that participated in PISA 2012, namely Hong Kong (China), Shanghai (China) and Macao (China), all of which are classified by the World Bank as high-income economies.

PISA 2012 assessed the competencies of 15-year-olds in reading, mathematics and science (with a focus on mathematics). The PISA 2012 data come from three sources: direct assessments of student performance, surveys of students, and surveys of schools. As PISA 2012 did not survey teachers, the information to characterise schools comes from the reports from school administrators about the physical academic and social characteristics of their schools, and from the reports from students about their schools, which are aggregated to the school level.

Three methodologies

The findings presented in this section come from original analyses based on three sets of models: *i)* multi-level analyses; *ii)* Oaxaca-Blinder counterfactual decomposition techniques (Blinder, 1973; Oaxaca, 1973); and *iii)* logistic regressions. Since the focus of PISA 2012 was on mathematics literacy, mathematics performance is the outcome used in these analyses, which use all five plausible values as estimated for each student. Each methodology is further described in the following sections.

Multi-level analyses for 18 middle-income countries

The multi-level models are specified as two-level regression models (student-level and school-level) within each country, with normally distributed residuals and maximum likelihood estimation. To account for the differences in sampling probabilities, both school and student weights are applied for all models, following standard OECD procedures (OECD, 2014). Models are estimated using Stata software.

The multi-level analysis began with an examination of each student-level variable separately and then in combination with other student-level variables; those variables most consistently related to mathematics performance were retained. Summary statistics are presented in Annex B. Various school-level factors were then examined to explore their additional contribution to student performance. This section discusses the results from the final regression in greater detail (Annex B, Model 12), but, for comparison, a number of different models are also included in Annex B and referred to in this section.

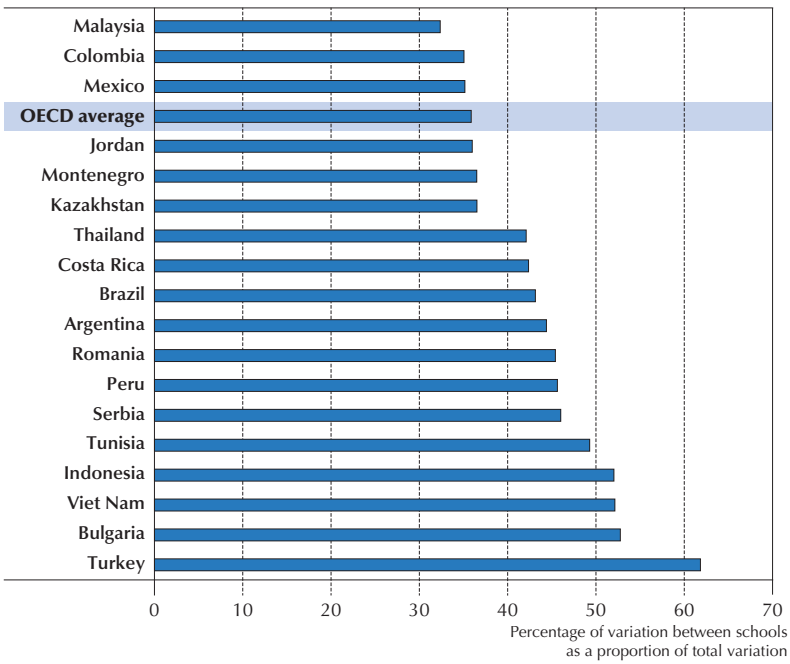
Variance in performance at two levels

The multi-level analysis begins with an analysis of the total variance in performance to be explained by student-level and school-level factors (Annex B, Model 1). These results can be presented as: *i)* the absolute size of the within-country performance variances; *ii)* the absolute size of the school-level and student-level performance variances; and *iii)* the share of the total variance in performance (between school variance plus within-school variance) that is due to variance between schools. This latter measure, the intra-class correlation, is referred to as "academic segregation" in the OECD documents.

Most middle-income countries exhibit more academic segregation compared to OECD countries. On average in OECD countries, 37% of the total variance in performance comes from "between school" variance in performance and 63% comes from "within-school" variance. This is not the pattern in most of the middle-income countries. In 11 of the 18 middle-income countries, the percentage of total variance accounted for by between-school variance is approximately 10 percentage points higher than for the OECD countries. In four of these countries, Bulgaria, Indonesia, Turkey and Viet Nam, the difference is greater than 15 percentage points. In some countries, the between-school share falls below that of OECD countries, but only by 1 or 2 percentage points (Figure 6.4).

Figure 6.4

Between school variance in mathematics performance, by country, PISA 2012



Source: Author's analysis; OECD (2013a), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201132-en (Table II.2.8a).

StatLink http://dx.doi.org/10.1787/888933293930

Education systems that exhibit a low share of total variation due to between-school variation indicate that student performance is likely to be the same regardless of which school they attend, but that students in the same school are performing at different levels. Education systems that exhibit a high share of total variance due to between-school variation indicate either that some schools are more "effective" than others, or that students are sorted into schools on the basis of their performance, and hence are performing at similar levels within these schools.

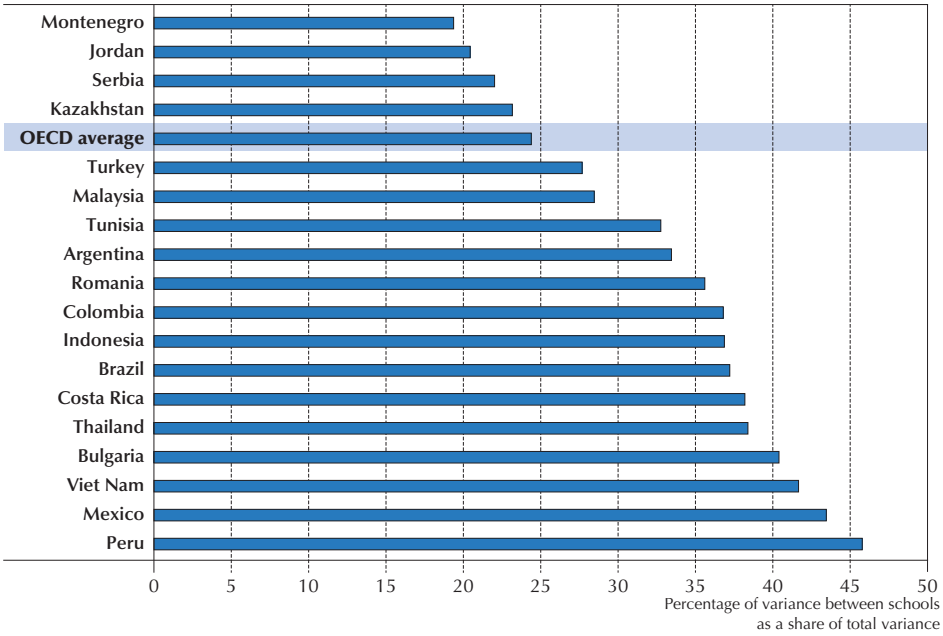
Variance in socio-economic status at two levels

Most middle-income countries exhibit greater social segregation compared with OECD countries. As with variance in performance, variance in the socio-economic status of students can be examined at two levels. The ratio of the between-school school variance in socio-economic status to the total variance in socio-economic status is defined as the "intra-class correlation" of this measure, which is called "social segregation."

On average in OECD countries, 24% of the variance in socio-economic status among students comes from "between school" variance in socio-economic status and 76% comes from "within school" variance. This is not the pattern for the middle-income countries in the PISA 2012 group (Figure 6.5). In 14 of these 18 countries, a much higher share of the variance in socio-economic student background occurs between schools compared to the OECD average, suggesting that the socio-economic background of students may be more of a determining factor in which school they attend. Unlike academic segregation, which is often a consequence of explicit educational policy, social segregation is likely to be a consequence of conscious or unconscious choices made by families, who may live in socio-economically homogeneous school "catchment" areas or may choose to enrol their children in private schools.

Figure 6.5

Between school variance in socio-economics status in 18 middle-income countries, PISA 2012



Source: OECD (2013a), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201132-en (Figure II.5.1b).

StatLink http://dx.doi.org/10.1787/888933293948

Social segregation is higher than the OECD average in all countries in Latin America, and in a few countries – Peru and Mexico, in particular – approximately 20 percentage points more of the variance in students' socio-economic background comes from between-school differences. This strongly suggests that students within the same school have similar socio-economic backgrounds, and that students having different types of socio-economic backgrounds do not attend the same schools. This has implications for how resources are distributed to schools and what policies may be needed to improve their quality.

A much higher share of both performance variation and socio-economic status variation occurs between schools in these middle-income countries than occurs between schools in OECD countries on average. However countries differ in the degree to which their schools are both academically and socially segregated. Eight countries exhibit both high academic segregation and high social segregation (Brazil, Bulgaria, Costa Rica, Indonesia, Peru, Romania, Thailand, Viet Nam, and four countries exhibit both low academic segregation and low social segregation Jordan, Kazakhstan, Malaysia, Montenegro). Colombia and Mexico exhibit high social segregation and low academic segregation, while Argentina, Serbia, Tunisia and Turkey exhibit high academic segregation and low social segregation.

Student-level influences on performance

In the multi-level analyses reported in *PISA 2012 Results: Volume IV* (OECD, 2013c), the OECD examined the effects on mathematics performance of students' socio-economic background, gender, language and the school's urban-rural location. The new multi-level analyses summarised in this section added four more characteristics at the student level: student grade, student grade repetition, school track (vocational versus general) and student pre-primary school attendance. These four characteristics are not immutable background characteristics and can be affected by education system policies related to the provision of pre-primary education, the age of school entry, requirements for progressing through the various grade levels, and tracking policies.

The new multi-level analyses found that a student's socio-economic background, gender, home language and residence were all strongly correlated with their PISA 2012 mathematics performance (Annex B, Model 5), which is consistent with the analyses reported by the OECD.

Students from more **socio-economically advantaged backgrounds** scored higher than those from less advantaged backgrounds in all countries other than Turkey. Home support for learning, as indicated by the number of books in the home,[7] is associated with better performance (Annex C).

In 14 countries, boys outperformed girls in mathematics, but **gender** was unrelated to mathematics performance in Jordan, Kazakhstan, Malaysia and Thailand (Annex B, Model 2). When all family background and school-level variables were considered simultaneously, boys outperformed girls in all countries other than Malaysia and Thailand (Annex B, Model 12).

Speaking the **language of the assessment** at home was positively associated with mathematics performance in nine countries (Argentina, Brazil, Bulgaria, Colombia, Jordan, Mexico, Peru, Tunisia and Viet Nam) and negatively associated with performance in two countries, (Indonesia and Thailand) possibly reflecting the presence of high-performing minority-language groups in these countries (Annex B, Model 3). When all family background and school-level variables were considered simultaneously, the language spoken at home continued to be positively correlated with performance in five countries (Bulgaria, Colombia, Jordan, Mexico and Peru), but in three countries (Indonesia, Malaysia, Thailand) those who did not speak the language of instruction at home outperformed those who did (Annex B, Model 12).

Students attending **schools located in a village** (and by inference, students living in a village, i.e. **rural students**) performed less well than those attending town or city schools (urban students) (Annex B, Model 4). These results are largely consistent with those reported by OECD. However, when all family background and school-level variables were considered simultaneously, students in rural schools performed the same as students in schools in other communities in all countries other than Brazil and Jordan, where their performance was higher (Annex B, Model 12).

In addition to these student characteristics, the new analyses looked at the effects of grade level, grade repetition, school track and preschool experience (all considered concurrently in Annex B, Model 10).

The **grade level** in which a student is enrolled strongly affects his or her performance, with students enrolled in higher grades scoring significantly above those enrolled in lower grades; this is the case for all countries other than Montenegro and Romania. In 13 countries, most 15-year-olds are enrolled in grade 10 (the "modal" grade in these countries), whereas in five countries most 15-year-olds are enrolled in grade 9 (the "modal" grade in these countries), but in all countries, students enrolled in a grade higher than the modal grade out-perform students in the modal grade, whereas students enrolled in grades below the modal grade perform less well, after controlling for repetition. This is undoubtedly due to differences in the opportunity to learn some skills that are tested on the PISA assessment.

The importance of **grade repetition**[8] in all 18 middle-income countries is visible from the large differences in mathematics performance between students who have and students who have not repeated a grade. Students who have repeated a grade at any time scored 30 points or more below those who had not repeated in Bulgaria, Jordan, Mexico, Montenegro and Serbia. This suggests that in some countries, students who are performing less well are retained in lower grades, keeping them in the system rather than excluding them.

Students in **vocational tracks** scored significantly above students in general education in three countries (Colombia, Costa Rica and Mexico) and significantly below students in general education tracks in seven countries (Bulgaria, Brazil, Kazakhstan, Montenegro, Malaysia, Serbia and Thailand), suggesting that either tracking policies based on student performance were in place or that students in the two tracks did not have the same opportunity to learn (Annex B, Model 7). In four countries (Jordan, Peru, Romania and Tunisia), not enough students were enrolled in vocational tracks to permit analysis.

A student's socio-economic background may be strongly related to a student's experience of **pre-primary education**, particularly when it is not universal and must be privately obtained. Nonetheless, even when the socio-economic background of the student is taken into account, the PISA 2012 mathematics scores of students with pre-primary experience were higher than the scores of those lacking this experience, although the level of this effect was statistically significant in only about half of the middle-income countries.

School-level influences on performance

Consistent with the previous OECD analysis, the new multi-level analyses find that once the socio-economic composition of the school is taken into account, most school inputs – infrastructure, educational resources, teacher quality and instructional time – are only modestly related to student performance (Annex B, Model 12).[9] These are all closely related to the socio-economic makeup of the schools themselves.

The **school physical infrastructure**, as measured by the PISA index, was related to student performance in only three countries: more resources were related to higher performance in Montenegro and Thailand and lower performance in Romania. The availability of educational resources was positively related to student performance in three countries: Brazil, Costa Rica and Romania. None of the **teacher quality** measures, including the proportion of teachers with higher education (International Standard Classification of Education [ISCED] 5A), were related to student performance. More **instructional time**, however, was positively associated with higher performance in eight countries (Argentina, Bulgaria, Indonesia, Kazakhstan, Mexico, Malaysia, Romania and Turkey).

The OECD analyses examined three **school climate** measures (disciplinary climate, student-teacher relationships, shared values) and the new multi-level analyses found that the disciplinary climate was the sole aspect of school climate that significantly contributed to an improvement of learning outcomes in middle-income countries. It is also one of the strongest and most consistent predictors of learning at the school level. Students in schools with a better disciplinary climate significantly outperformed students in schools with a less positive disciplinary climate in 13 countries. A school's average level of student absenteeism was strongly and negatively related to performance.

The new multi-level regression analyses show that **opportunity to learn** (i.e. exposure to formal mathematics) was associated with higher mathematics performance in all middle-income countries, but may also reflect a student's grade level and national study programme (Annex B, Model 11).

Institutions

Private education and tracking are strongly related to student performance in many middle-income countries. In eleven middle-income countries, over 5% of students were enrolled in a private school (Annex B, Descriptive Statistics). Students enrolled in **private schools** outperformed those enrolled in public schools in eight countries (Annex B, Model 8), but these performance differences were sharply reduced once school and student socio-economic status were taken into account (Annex B, Model 8a). In relation to **tracking**, four middle-income countries reported a very high share of students in vocational programmes: Serbia (74%), Montenegro (66%), Bulgaria (41%) and Turkey (38%). All of these were significantly higher shares than in OECD countries (OECD, 2013c, Table IV.2.6).

Oaxaca-Blinder decomposition of private school and tracking effects

Two systemic differences that exist among countries are explored in this analysis: *i)* the degree of stratification of their educational systems through tracking; and *ii)* the nature of selection for different schools, particularly private schools. Highly stratified systems, such as those in Bulgaria, Montenegro and Serbia, practice early selection based on prior academic ability and group students into differentiated educational programmes (typically, vocational tracks versus general education tracks). In comprehensive systems, education is not differentiated by ability. However, private education offers competition to public education and access is largely determined by a family's economic status, particularly in Latin American countries.[10] Both types of differences are associated with large differences in student performance in some middle-income countries.

This section analyses the factors that account for these two difference in performance between general and vocational schools and between public and private schools using the decomposition method often referred to as "Oaxaca-Blinder" decomposition (Blinder, 1973; Oaxaca, 1973). The decomposition technique was originally used in sociology and demography, but was popularised in economics literature to decompose male/female earnings gaps. It has since been used to decompose learning gaps in education in Indonesia (Barrera-Osorio et al., 2011). In this section, the same method is used to examine the gap between student performance in vocational versus academic tracks, and between student performance in public and private schools.

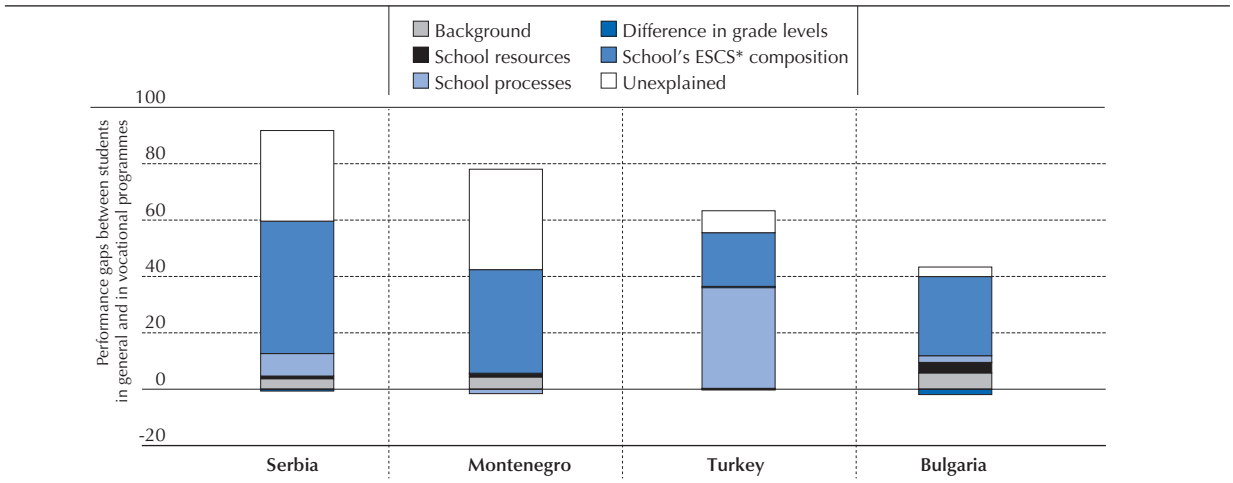
The method, as applied in this section, divides the mathematics performance score differential between two groups into one part that is "explained" by observable group differences in family, student and school characteristics, and a residual part that cannot be accounted for. This "unexplained" part often subsumes the effects of group differences in unobserved predictors (for a detailed description of the technique used see Jann, 2008).

Accounting for the vocational-academic gap

Early tracking appears to separate students into general education programmes and vocational programmes based on their performance towards the end of the basic education cycle. Performance gaps between students enrolled in general programmes and those enrolled in vocational programmes are found among 15-year-olds participating in PISA and are the largest in Serbia (91 points higher in general schools) and Montenegro (76 points), followed by Turkey (63 points); they are much less pronounced in Bulgaria (41 points). A decomposition of these gaps shows that in Bulgaria, Montenegro and Serbia, the largest share of the performance gaps between students in these two types of programme can be accounted for by differences in the socio-economic composition of students in general versus vocational schools; this is not the case in Turkey (Figure 6.6). In Turkey, the largest share of the performance gap can be accounted for by differences in school processes (disciplinary climate and instructional time), followed by differences in the socio-economic composition of the schools. In both Serbia and Montenegro, however, much of the performance gap remains unaccounted for.

Figure 6.6

Decomposition of mathematics performance gap between general and vocational programmes in 4 middle-income countries, PISA 2012



* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: Authors' original analyses of PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933293951

Accounting for the public-private gap

Nine of the 18 middle-income countries have a high share of students enrolled in private schools: all the Latin American countries (9%-32%), Indonesia (41% of students), Jordan (17%), and Thailand (17%). The following countries have fewer than 2% of students enrolled in private schools: Bulgaria, Montenegro, Romania, Serbia, Tunisia and Turkey.

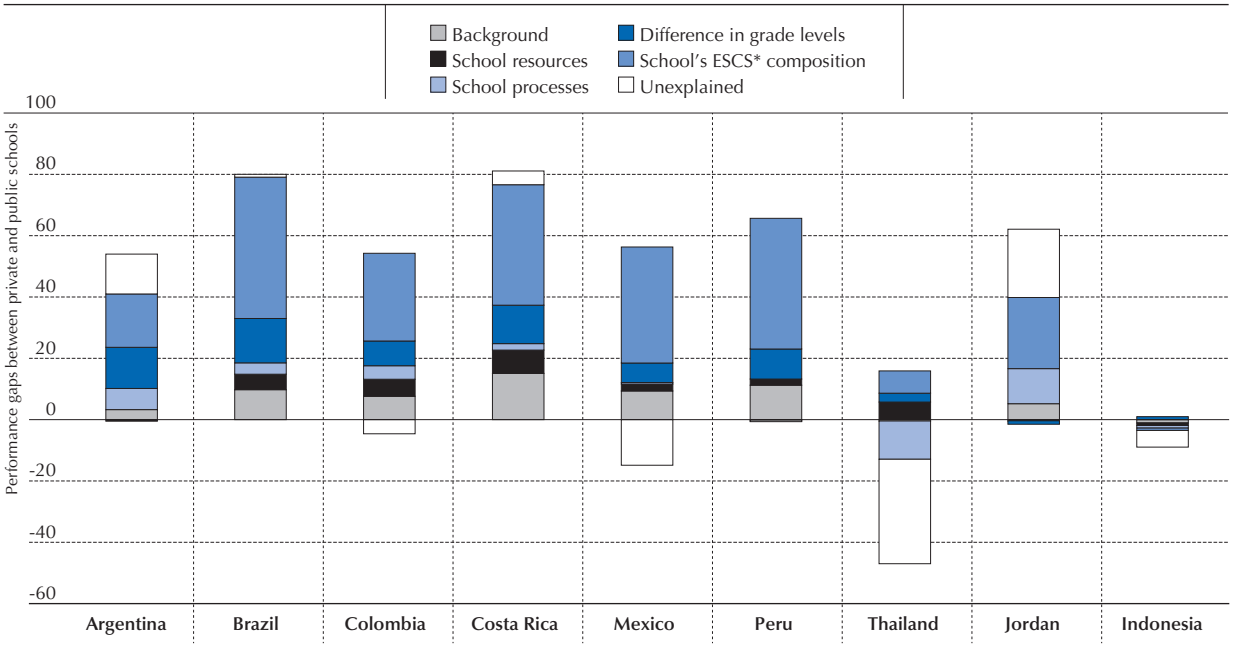
Gaps in performance between private and public schools, with students in private schools outperforming those in public schools, are large in Malaysia (87 points), Brazil (80 points), Costa Rica (78 points), Peru (65 points), Jordan (60 points), Argentina (58 points), Colombia (50 points), and Mexico (42 points). In Thailand, students in private schools underperform students in public schools (Annex B Model 8).

Performance gaps between public and private schools can be largely explained by: *i)* students' socio-economic background; *ii)* the school's socio-economic composition; and *iii)* the high share of students in a grade below the modal grade for 15-year-olds. Little is left unexplained in most of the Latin American countries, although large amounts are left unexplained in Thailand and Jordan (Figure 6.7). This suggests that social segregation is an important factor in explaining performance differences across students, particularly for Brazil, Colombia, Costa Rica, Mexico and Peru.

Students from disadvantaged backgrounds are often doubly disadvantaged: not only do they come from a home environment that may be less encouraging of school learning, but the schools they attend may also be less advantaged. For all countries other than Thailand and Indonesia, most of the performance difference between public school students and private schools students can be attributed to differences between the students that attend such schools. Similarly, most differences between the performance of students in vocational schools and those in general secondary schools can be attributed to differences in the socio-economic status of students attending these two different types of schools.

Figure 6.7

Decomposition of mathematics performance gap between private and public schools in 9 middle-income countries, PISA 2012



* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: Authors' original analyses of PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933293963

Odds of being in a disadvantaged school

The analysis of PISA data undertaken in this chapter has underscored the importance of both a student's own socio-economic background and the average socio-economic background of a school's students as strong correlates of performance. The analyses have also suggested that when student and school socio-economic background are taken into account, most other school inputs are considerably less related to student performance.

The final set of analyses carried out for this report examines the odds of being enrolled in an advantaged school. Advantaged schools are defined as having one or more of the following: an adequate school building and grounds, sufficient classrooms, electricity and climate control, qualified mathematics teachers, adequate instructional materials, and adequate library materials. Each of these advantages are examined separately, and the questions are: do students with a higher socio-economic background have higher odds of attending a school that has one or more of these advantages, compared to students from a lower socio-economic background? And are these odds different in countries with relatively higher shares of students in vocational schools or private schools? No countries have both a high share of students in vocational programmes and a high share in private schools. Table 6.8 summarises these odds; details are provided in Annex B.

Table 6.8

Odds of a student being enrolled in schools with adequate inputs, selected countries, PISA 2012

School resource advantage	Countries where students with higher socio-economic background have higher odds of attending a school with resource advantage	
	Countries with more than 35% of students in vocational programmes (Bulgaria, Montenegro, Serbia, Turkey)	**Countries with more than 9% of students in private schools** (Argentina, Brazil, Colombia, Costa Rica, Indonesia, Mexico, Peru, Thailand)
Qualified mathematics teachers	Bulgaria, Montenegro, Turkey	Brazil, Indonesia Jordan, Mexico, Peru
Instructional materials	Bulgaria, Montenegro	Brazil, Colombia, Costa Rica, Indonesia, Mexico, Peru
Library materials	Bulgaria, Montenegro, Turkey	Brazil, Colombia, Costa Rica, Indonesia, Jordan, Mexico, Peru
School buildings and grounds	Bulgaria, Montenegro, Turkey	Argentina, Brazil, Colombia, Costa Rica, Indonesia, Jordan, Mexico, Peru
Heating/cooling and lighting	Montenegro, Turkey	Argentina, Brazil, Costa Rica, Indonesia, Jordan, Mexico, Peru, Thailand
Instructional space	Bulgaria, Montenegro, Turkey	Argentina, Brazil, Costa Rica, Indonesia, Mexico, Peru

Source: Author's original analysis of PISA 2012 Database.

Students from higher socio-economic backgrounds in three of the four countries with early tracking, as indicated by the share of students in vocational programmes, are more likely to be attending schools that have higher quality educational resources (Bulgaria, Montenegro and Turkey). Similarly, students from socially-advantaged backgrounds in most of the countries with a substantial share of students enrolled in private schools are more likely to be attending schools that have higher quality educational resources (Argentina, Brazil, Colombia, Costa Rica, Indonesia, Jordan, Mexico and Peru). There are, however, exceptions: in Serbia, a country with a high share of students in vocational programmes, and in Thailand, a country with a high share of private schools, socially advantaged students are not more likely to be in schools that have higher quality educational resources.

Different mechanisms appear to lie behind the sorting of students into different schools in low- and middle-income countries. In some Latin American countries, social segregation and income inequality is reflected in the high social segregation between schools; ethnicity and rural residence are often associated with these differences. In the former socialist countries and Turkey, academic excellence is the sorting mechanism behind channelling children into different schools. As a consequence, these countries combine low levels of social segregation with high levels of academic segregation. Somewhere in between these two cases is Viet Nam, the best-performing country among the middle-income countries studied, which combines high academic segregation with relatively high social segregation between schools.[11]

CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES

Conclusions

PISA provides participating countries with considerable information that can enable them to identify issues related to education quality and equity within their own country. In particular:

1. PISA shows that the quality of human capital in middle-income countries, as measured by PISA assessments, is, with the exception of Viet Nam, lower than in high-income countries. In accordance with their economic development, all of these middle-income countries have considerably fewer financial resources invested in their education systems. The average socio-economic background of the students is also much lower than in the high-income countries.
2. PISA identifies educational inequalities that vary among countries, but that are associated with a student's gender, socio-economic status and home language (a proxy for ethnicity).
3. PISA demonstrates that socio-economically disadvantaged students frequently attend resource-disadvantaged schools in middle-income countries.

PISA also provides information about aspects of the education system overall that can be the topic of policy discussion. In particular, PISA identifies two important aspects of stratification in education: academic stratification and social stratification:

1. Some former socialist countries and Turkey, unlike most OECD countries, combine high academic segregation between schools with low social segregation within schools. The latter is a direct consequence of traditionally low socio-economic inequalities beyond the education system, and low residential segregation in these countries. In these systems, the socio-economic background of students may play a role in sorting students into different school tracks, but is less relevant once the students have been sorted.
2. At the other extreme are many Latin American countries, where socio-economic inequalities result in social stratification between schools but low academic segregation within schools. In these countries, a lack of academic selection enables wealth to be the sole mechanism for sorting children across schools. As a result, the socially advantaged have greater access to high levels of educational resources, often provided through elite private schools.

The new multi-level analysis of the PISA 2012 mathematics assessment identified four school-level characteristics that are associated with higher student performance in middle-income countries, after the effects of other background characteristics are statistically controlled. These are: learning and instructional time, school climate, engagement with and at school, and opportunities to learn.

At the same time, the PISA results to date may have less relevance for middle-income countries, for two reasons: *i)* the comparability across countries of the students assessed; and *ii)* the relevance of the PISA performance measures themselves for middle-income countries.

1. As mentioned above, in some countries a substantial proportion of the 15-year-olds enrolled in school are enrolled in grades considerably lower than the modal grade for their age and country: often in grades 7 or 8, rather than grades 9 or 10 as in OECD countries. In addition, out-of-school rates for lower secondary school children are high in many low- and middle-income countries, exceeding 10% in Argentina, Bulgaria, Jordan, Mexico, Thailand and Turkey. The rate is exceptionally high in Viet Nam, at 37%. This implies selection bias among the population eligible for the PISA survey in these countries. The combination of these two exclusion mechanisms, plus other minor exclusion criteria,[12] result in indices of coverage of the 15-year-old population as low as 50% in Costa Rica, 56% in Viet Nam, and 63% in Indonesia, compared with 96% for OECD countries. These relatively low levels of coverage limit the comparability of the results of middle-income countries with other countries.
2. The characteristics of the PISA mathematics scores in middle-income countries also differ somewhat from those in high-income OECD countries (for which the tests were originally designed), which have a mean of about 500 and a standard deviation of about 100. In the middle-income countries, scores on the PISA 2012 mathematics assessment, while normally distributed, have much lower means and a smaller standard deviation. In many countries, well over 50% of students achieve scores lower than 420 on the scale (at Level 1 or below in mathematics proficiency), compared with an OECD average of 23%.

Finally, the analyses carried out in this chapter point to several areas that could be improved for PISA survey instruments to better represent the situation in middle-income countries:

1. While PISA aims to collect information on the quality of teachers and teaching, a great deal of essential information about these important determinants of learning in developing countries is missing from PISA. Moreover, the quality of information that is provided is hampered by imprecision: principals assess the general quality of the teachers in their schools, rather than the quality of individual teachers. A principal's bias could affect this measure; moreover, this measure does not provide information about individual teachers. Specifically, more precise information is needed about teachers' knowledge of the subject content and about the level of their pedagogical skills. More information is also needed on instructional processes, particularly about how class time is spent, for example in independent activities, such as working in workbooks, versus small group activity and whole-class teacher-centred instruction (Willms and Tramonte, 2015). The OECD could assist partner countries in improving the collection of information on this very important aspect of the learning environment as an aid to development projects.
2. There is much evidence that variations in the amount of educational resources available to students in schools affect learning in developing countries, yet this relationship is not significant in the PISA study. The PISA survey instruments emphasise features of schools that may be less important in lower-middle-income countries, such as availability of computers and the Internet. Moreover, these school-level resources are often linked to the socio-economic composition of schools, particularly in Latin America, where the difference in resource allocation between advantaged and disadvantaged schools is very high.

Implications

There are several implications for the future utilisation of PISA for greater application to middle-income countries, for both middle-income countries and the OECD.

Implications for middle-income countries

In order to improve the relevance of PISA for middle-income countries, countries are encouraged to participate more actively in the development of instruments and the completion of analyses relating to policy issues of concern in their countries.

Implications for the OECD

As has been anticipated by the PISA for Development initiative, the OECD will need to acknowledge the differences between the countries currently and historically participating in PISA, and the low- and middle-income countries that are not currently participating in PISA, by:

1. Adapting the cognitive instruments to better capture performance differences at the lower end of the performance scale. An analysis of the needs in this respect is currently underway by Adams and Cresswell (2014).
2. Adapting the questionnaires to better measures institutions and inputs that are relevant to these countries. An analysis of the questionnaires is currently underway by Willms and Tramonte (2015).

The OECD will also need to consider how to revise some of the regular OECD analyses of PISA to focus on middle-income country participants, rather than grouping all non-OECD partner countries/economies into a single group by:

1. Analysing results separately by income group.
2. Analysing results separately by world region.

Notes

1. The following middle-income countries/economies are not included, for the following reasons: *i)* Albania lacked information on student socio-economic status; *ii)* China's data referred to Shanghai only; and *iii)* Hungary was a high-income country from 2009 to 2013.

2. This chapter is based on a background paper prepared by Tijana Prokic-Breuer.

3. OECD measures mathematics proficiency on a scale having a mean value of 500 and a standard deviation of 100. Six proficiency levels are reported: Level 6 (669 points or above), Level 5 (607-668 points), Level 4 (545-606 points), Level 3 (482-544), Level 2 (420-481), Level 1 (358-419).

4. Issues of endogeneity affect this measure, since students who are attending their lessons may both recall more exposure and have higher mathematics performance.

5. The OECD provides many reasons why school-level variables may not be associated with student achievement; nevertheless, the OECD does carry out these analyses.

6. Reasons for exclusion were: *i)* Albania lacked information on student socio-economic status; and *ii)* Hungary was a high-income country from 2009 to 2013.

7. See Carnoy and Rothstein (2013) for a discussion of this variable.

8. Controlling for grade of enrolment; repetition is dummy indicating whether or not the student has reported having repeated a grade in either primary, secondary or upper secondary school.

9. To avoid "overcontrolling" the models presented in Annex B contain fewer school-level variables than the ones presented in OECD, 2014. Moreover, the variables in the new analyses were introduced carefully into the basic model and tested for their independent effects before incorporating them into the full model.

10. In some Latin American countries, voucher systems open the opportunity for disadvantaged students to afford private schooling.

11. Viet Nam's high performance in PISA 2012 has attracted considerable attention. However, in Viet Nam, only about 55% of 15-year-olds are in school in eligible PISA grades, suggesting that the average performance would be lower if a representative sample of 15-year-olds were selected for assessment.

12. Details can be found in OECD, 2014.

References

Adams, R. and **J. Cresswell** (2014), *PISA for Development Technical Strand 2: Enhancement of PISA Cognitive Instruments*, Australian Council for Education Research.

Akiba, M., G. LeTendre and **J. Scribner** (2007), "Teacher Quality, Opportunity Gap, and National Achievement in 46 Countries", *Educational Researcher,* Vol. 36, No. 7, pp. 369-387, American Educational Research Association, Washington, DC.

Austin, B., O. Adesope, B. French, C. Gotch, J. Bélanger and K. Kubacka (2015), "Examining School Context and its Influence on Teachers: Linking Talis 2013 with PISA 2012 Student Data", *OECD Education Working Papers*, No. 115, OECD Publishing, Paris, http://dx.doi.org/10.1787/5js3f5fgkns4-en.

Baker, D., B. Goesling and **G. LeTendre** (2002), "Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the 'Heyneman-Loxley Effect' on Mathematics and Science Achievement", *Comparative Education Review*, Vol. 46, No. 3, pp. 291-312, University of Chicago Press, Chicago.

Barrera-Osorio, F., V. Garcia Moreno, H. Patrinos and **E. Porta** (2011), "Using the Oaxaca-Blinder decomposition technique to analyze learning outcomes changes over time: An application to Indonesia's results in PISA mathematics", *Regional and Sectoral Economic Studies,* Euro-American Association of Economic Development, Vol. 11, No. 3.

Berlinski, S., S. Galiani and **P. Gertler** (2009), "The Effect of Pre-primary Education on Primary School Performance", *Journal of Public Economics*, Vol. 93, No.1, pp. 219-234, Elsevier, Amsterdam.

Blinder, A.S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", *The Journal of Human Resources*, Vol. 8, pp. 436-455, University of Wisconsin Press, Madison.

Carnoy, M. and **R. Rothstein** (2013), "What Do International Tests Really Show About US Student Performance", Economic Policy Institute, www.epi.org/publication/us-student-performance-testing/.

Engle, P., M. Black, J. Behrman, M. de Mello, P. Gertler, L. Kapiriri and **M. Young** (2007), "Child Development in Developing Countries: Strategies to Avoid the Loss of Developmental Potential in More Than 200 Million Children in the Developing World", *Lancet*, Vol. 369, No. 9557, pp. 229-242, Elsevier, Amsterdam.

Evans, D. and **A. Popova** (2015), "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews", *Policy Research Working Papers,* WPS7203, World Bank, Washington DC.

Fuchs, T. and **L. Wößmann** (2007), "What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data", *Empirical Economics*, Vol. 32, No. 2-3, pp. 433-464, Springer.

Fuller, B. (1987), "What School Factors Raise Achievement in the Third World?", *Review of Educational Research*, Vol. 57, No. 3, pp. 255-292, Sage, Los Angeles.

Glewwe, P., E. Hanushek, S. Humpage and **R. Ravia** (2014), "School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010", in P. Glewwe (ed.), *Educational Policy in Developing Countries*, University of Chicago Press, Chicago.

Hanushek, E. (1997), "Assessing the Effects of School Resources on Student Performance: An Update", *Educational Evaluation and Policy Analysis*, Vol. 19, No. 2, pp. 141-164, Sage, Los Angele.

Hanushek, E. and **L. Wößmann** (2014), "Institutional Structures of the Education System and Student Achievement: A Review of Cross-Country Economic Research", in R. Strietholt, W. Bos, J. Gustafsson, Rosen and M. Munster (eds.), *Educational Policy Evaluation through International Comparative Assessments*, pp. 145-175, Waxman, Germany.

Hanushek, E., S. Link and **L. Wößmann** (2013), "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA", *Journal of Development Economics,* Vol. 104, pp. 212-232, Elseview, Amsterdam, National Bureau of Economic Research, Cambridge, MA.

Hattie, J. (2013). *Visible Learning: A Synthesis of over 800 Meta-analyses Relating to Achievement*, Routledge, New York.

Heyneman, S. and **W. Loxley** (1983), "The Effect of Primary-School Quality on Academic Achievement Across Twenty-Nine High- and Low-Income Countries", *American Journal of Sociology*, pp. 1162-1194, University of Chicago Press, Chicago.

Jann, B. (2008), "The Blinder-Oaxaca Decomposition for Linear Regression Models", in *The Stata Journal*, Vol. 8, No. 4, pp. 453-479, StataCorp.

Kremer, M., C. Brannen and **R. Glennerster** (2013), "The Challenge of Education and Learning in the Developing World", *Science*, Vol. 340/6130, pp. 297-300, American Academy for the Advancement of Science, Washington, DC.

Lavy, V. (2010), "Do Differences in Schools' Instruction Time Explain International Achievement Gaps in Math, Science and Reading? Reading Evidence from Developed and Developing Countries", *NBER Working Papers* 16227, National Bureau of Economic Research, Cambridge MA.

Lewis, M. and **M. Lockheed** (2007), *Inexcusable Absence: Why 60 Million GirlsStill Aren't in School and What to do About It*, Center for Global Development, Washington, DC.

Lockheed, M. and **A. Verspoor** (1990), *Improving Primary Education in Developing Countries*, Oxford University Press, New York.

Lockheed, M. and **E. Hanushek** (1988), "Improving Educational Efficiency in Developing Countries: What do we know?", *Compare* Vol. 18, No. 1, pp. 21-38, Taylor and Francis, London.

Long, D. (2013), "Cross-national Educational Inequalities and Opportunities to Learn: Conflicting Views of Instructional Time", *Educational Policy*, Sage, Los Angeles.

McEwan, P. (2014), "Improving Learning in Primary Schools of Developing Countries: A Meta-analysis of Randomized Experiments", *Review of Educational Research*, Vol. 85, No. 3, pp. 353-394, Sage, New York.

Murnane, R. and **A. Ganimian** (2014), "Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations," *NBER Working Papers*, No. 20284, National Bureau of Economic Research, Cambridge, MA.

Oaxaca, R. (1973), "Male-female Wage Differentials in Urban Labor Markets", in *International Economic Review*, Vol. 14, pp. 693-709, John Wiley and Sons.

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

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

OECD (2013b), *PISA 2012 Results: Ready to Learn (Volume III): Students' Engagement, Drive and Self-Beliefs*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201170-en.

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

Raudenbush, S. and **J.D. Willms, Eds.,** (1991), *Schools, Classrooms and Pupils: International Studies of Schooling from a Multilevel Perspective,* Academic Press, San Diego.

Riddell, A. (2008), *Factors Influencing Educational Quality and Effectiveness in Developing Countries: A Review of Research*, GTZ, Eschborn, Germany.

Riddell, A. (1989), "An Alternative Approach to the Study of School Effectiveness in the Third World Countries", *Comparative Education Review*, Vol. 33, No. 4, pp. 481-497, University of Chicago Press, Chicago.

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

Schütz, G., H. Ursprung and **L. Wößmann** (2008), Education Policy and Equality of Opportunity, Kyklos, Vol. 61, pp. 279-308, http://dx.doi.org/10.1111/j.1467-6435.2008.00402.x.

Simmons, J. and **L. Alexander** (1978), "The Determinants of School Achievement in Developing Countries: A Review of the Research", *Economic Development and Cultural Change*, pp. 341-357, University of Chicago Press, Chicago.

UNESCO (2015), *Education for all 2000-2015: Achievements and Challenges*, UNESCO, Paris.

Walberg, H. J. (1984). "Improving the productivity of America's schools." *Educational Leadership,* Vol. 41, No. 8, pp. 19-27, Association for Supervision and Curriculum Development, Alexandria, VA.

Wang, M., G. Haertel and **H. Walberg** (1990), "What influences learning? A content analysis of review literature", *The Journal of Educational Research*, Vol. 84, No. 1, pp. 30-43, Taylor and Francis, New York.

Willms, J. and **L. Tramonte** (2015), "Towards the development of contextual questionnaires for the PISA for development study", *OECD Education Working Papers,* No. 118, OECD Publishing, Paris, http://dx.doi.org/10.1787/5js1kv8crsjf-en.

Wößmann, L. (2009), "International Evidence on School Tracking: A review", *Journal for Institutional Comparisons*, Vol. 7, No. 1, pp. 26-34.

Wößmann, L. (2007), "International Evidence on School Competition, Autonomy, and Accountability: A Review", *Peabody Journal of Education*, Vol. 82, No. 2-3, pp. 473-497, Taylor and Francis, New York.

Wößmann, L. (2005), "The Effect Heterogeneity of Central Examinations: Evidence from TIMSS, TIMSS-Repeat and PISA", *Education Economics*, Vol. 13, No. 2, pp. 143-169, Taylor and Francis, New York.

Wößmann, L. (2003), "Central Exit Exams and Student Achievement: International Evidence", in P.E. Peterson and M.R. West (eds.), *No Child Left Behind? The Politics and Practice of School Accountability*, pp. 292-323, Brookings Institution Press, Washington, DC.

PISA FOR DEVELOPMENT

Annex A

Selected tables from PISA 2012 results, Volumes I-IV

Analyses in this report were based on PISA 2012 data, using the following tables. The tables are accessible on line following the links indicated below.

Volume 1

PISA 2012 Results: What Students Know and Can Do: Student Performance in Mathematics, Reading and Science **(OECD, 2014)**

- PISA target populations and samples, Table A2.1, http://dx.doi.org/10.1787/888932937092.
- Percentage of students at each proficiency level in mathematics, Table I.2.1a, http://dx.doi.org/10.1787/888932935667.
- Percentage of students below Level 2 and at Level 5 or above in mathematics in PISA 2003 through 2012, Table I.2.1b, http://dx.doi.org/10.1787/888932935667.
- Mean score, variation and gender differences in student performance in mathematics, Table I.2.3a, http://dx.doi.org/10.1787/888932935667.
- Percentage of students at each proficiency level in reading, by gender, Table I.4.2a, http://dx.doi.org/10.1787/888932935705.

Volume 2

PISA 2012 Results: Excellence through Equity: Giving Every Student the Chance to Succeed **(OECD, 2013a)**

- Variation in mathematics performance, Table II.2.8a, http://dx.doi.org/10.1787/888932964908.
- Students' socio-economic status, Table II.2.13a, http://dx.doi.org/10.1787/888932964908.
- Relationship between performance in mathematics, reading and science, and socio-economic status, Table II.2.1, http://dx.doi.org/10.1787/888932964908.
- Relationship between mathematics performance and socio-economic status, between and within schools, Table II.2.9a, http://dx.doi.org/10.1787/888932964908.
- Equity in opportunity to learn: Formal mathematics, Table II.4.1, http://dx.doi.org/10.1787/888932964946.
- Inequity in access to educational resources: Disciplinary climate, Table II.4.10, http://dx.doi.org/10.1787/888932964946.

Volume 3

PISA 2012 Results: Ready to Learn: Students' Engagement, Drive and Self-Beliefs **(OECD, 2013b)**

- Mathematics performance, by the number of times students arrived late for school in the two weeks prior to the PISA test, Table III.2.1a, http://dx.doi.org/10.1787/888932963920.
- Association between arriving late for school and mathematics performance, by performance level, Table III.2.1.c, http://dx.doi.org/10.1787/888932963920.
- Mathematics performance, by the number of times students skipped a day of school in the two weeks prior to the PISA test, Table III.2.2b, http://dx.doi.org/10.1787/888932963920.
- Association between skipping classes or days of school and mathematics performance, by performance level, Table III.2.2c, http://dx.doi.org/10.1787/888932963920.

Volume 4

***PISA 2012 Results: What Makes a School Successful? Resources, Policies and Practices* (OECD, 2013c)**

- Grade repetition,Table IV.2.2, http://dx.doi.org/10.1787/888932957422.
- Student grade level, Table IV.2.4, http://dx.doi.org/10.1787/888932957422.
- Preschool attendance, Table IV.3.33, http://dx.doi.org/10.1787/888932957460.
- Students' learning time in school, Table IV.3.21, http://dx.doi.org/10.1787/888932957460.
- Students' learning time in school, by school features, Table IV.3.22, http://dx.doi.org/10.1787/888932957460.
- Quality of physical infrastructure, by school features, Table IV.3.15, http://dx.doi.org/10.1787/888932957460.
- Index of quality of schools' educational resources, by school features, Table IV.3.17, http://dx.doi.org/10.1787/888932957460.
- Pre-service teacher training requirements in public institutions, Table IV.3.4, http://dx.doi.org/10.1787/888932957460.
- Student-teacher ratios, by school features, Table IV.3.9, http://dx.doi.org/10.1787/888932957460.
- Composition and qualifications of teaching staff, Table IV.3.6, http://dx.doi.org/10.1787/888932957460.
- Teacher professional development, Table IV.3.13, http://dx.doi.org/10.1787/888932957460.

References

OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201118-en.

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

OECD (2013b), *PISA 2012 Results: Ready to Learn (Volume III): Students' Engagement, Drive and Self-Beliefs,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264201170-en.

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

PISA FOR DEVELOPMENT

Annex B

Results of multi-level analyses of PISA 2012 mathematics in 18 middle-income countries

ANNEX B: RESULTS OF MULTI-LEVEL ANALYSES OF PISA 2012 MATHEMATICS IN 18 MIDDLE-INCOME COUNTRIES

ISO codes used for the 18 middle-income countries

ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Argentina	Bulgaria	Brazil	Colombia	Costa Rica	Indonesia	Jordan	Kazakhstan	Mexico

MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Montenegro	Malaysia	Peru	Romania	Serbia	Thailand	Tunisia	Turkey	Viet Nam

[Part 1/1]

Table B.1 **Model 1: Empty model**

Empty model	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Constant	388.8***	439.4***	389.3***	377.0***	407.2***	375.2***	385.7***	431.5***	413.9***
	(100.88)	(84.26)	(144.72)	(112.26)	(106.06)	(81.98)	(103.06)	(126.54)	(246.97)
Between school variance	2 597.5***	4 646.8***	2 716.6***	1 951.8***	1 983.7***	2 665.3***	2 165.0***	1 860.3***	1 938.5***
	(84.60)	(82.93)	(76.68)	(59.86)	(49.59)	(47.88)	(43.89)	(58.29)	(135.29)
Within school variance	3 253.2***	4 160.3***	3 434.3***	3 617.7***	2 699.4***	2 456.3***	3 852.1***	3 234.5***	3 578.2***
	(300.07)	(321.05)	(357.86)	(266.62)	(261.20)	(255.97)	(325.30)	(301.74)	(540.02)
Number of observations	5 908	5 282	19 204	9 073	4 602	5 622	7 038	5 808	33 806

Empty model	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Constant	409.8***	420.5***	368.7***	444.6***	449.2***	426.5***	388.0***	448.2***	511.7***
	(42.58)	(107.59)	(88.89)	(101.92)	(85.34)	(104.56)	(83.34)	(74.14)	(99.29)
Between school variance	2 485.2***	2 128.8***	3 243.4***	2 986.6***	3 775.7***	2 865.2***	3 017.0***	5 108.0***	3 823.6***
	(26.04)	(57.37)	(71.55)	(72.13)	(64.61)	(64.45)	(49.51)	(67.16)	(66.04)
Within school variance	4 323.6***	4 449.3***	3 864.6***	3 588.7***	4 430.6***	3 941.1***	3 104.2***	3 172.7***	3 507.6***
	(206.31)	(259.85)	(310.45)	(299.45)	(269.47)	(303.13)	(302.83)	(268.99)	(260.27)
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295624

[Part 1/1]

Table B.2 **Model 2: Mathematics performance and gender**

Individual features	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Gender (ref. male)	-16.09***	-15.19***	-20.73***	-25.90***	-24.83***	-7.028***	-2.594	-2.132	-18.01***
	(-6.80)	(-7.41)	(-14.23)	(-12.03)	(-14.36)	(-3.84)	(-0.32)	(-1.20)	(-17.92)
Constant	397.1***	446.7***	400.0***	390.7***	420.4***	378.6***	387.0***	432.5***	423.0***
	(96.40)	(81.06)	(138.47)	(104.57)	(101.60)	(79.44)	(56.24)	(118.52)	(235.52)
Between school variance	2 609.1***	4 767.3***	2 752.2***	1 940.1***	1 999.1***	2 672.6***	2 195.5***	1 862.6***	1 971.6***
	(84.82)	(84.18)	(77.60)	(60.57)	(50.25)	(47.94)	(41.19)	(58.21)	(136.66)
Between school variance	3 191.1***	4 107.3***	3 325.0***	3 461.4***	2 543.9***	2 445.0***	3 850.2***	3 233.5***	3 496.0***
	(300.18)	(318.17)	(366.09)	(261.88)	(257.90)	(257.82)	(324.26)	(302.01)	(531.65)
Number of observations	5 908	5 282	19 204	9 073	4 602	5 622	7 038	5 803	33 806

Individual features	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Gender (ref. male)	-12.86***	3.097	-27.37***	-14.47***	-25.14***	-1.407	-24.17***	-22.15***	-25.14***
	(-3.65)	(1.44)	(-13.54)	(-5.87)	(-9.65)	(-0.70)	(-12.67)	(-11.20)	(-11.68)
Constant	416.2***	418.9***	382.7***	452.0***	461.8***	427.3***	400.9***	459.2***	525.1***
	(39.69)	(98.31)	(85.78)	(97.49)	(80.14)	(95.91)	(82.12)	(70.76)	(92.75)
Between school variance	2 559.1***	2 120.9***	3 338.2***	3 058.3***	3 962.3***	2 876.4***	3 130.3***	5 237.0***	4 018.6***
	(26.59)	(57.40)	(72.01)	(73.93)	(66.36)	(64.41)	(50.46)	(67.77)	(66.74)
Between school variance	4 284.5***	4 447.5***	3 702.4***	3 537.4***	4 285.9***	3 940.2***	2 955.1***	3 072.8***	3 345.9***
	(193.65)	(260.96)	(303.36)	(300.49)	(269.03)	(303.37)	(294.25)	(257.76)	(257.93)
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295630

[Part 1/1]

Table B.3 **Model 3: Language disparities**

Individual features	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Native speaker	12.36**	27.41***	16.12***	26.76***	5.606	-10.16***	23.59***	1.571	22.64***
	(2.73)	(6.47)	(4.14)	(4.48)	(0.91)	(-4.48)	(5.19)	(0.34)	(7.77)
Constant	377.0***	415.4***	373.9***	350.9***	401.8***	379.3***	363.9***	430.0***	392.5***
	(66.25)	(67.56)	(78.47)	(50.00)	(56.09)	(81.84)	(59.43)	(79.20)	(124.75)
Between school variance	2 565.8***	4 300.6***	2 703.8***	1 938.2***	1 981.9***	2 724.8***	2 131.8***	1 864.2***	1 872.9***
	(85.06)	(80.56)	(76.50)	(60.18)	(49.54)	(47.02)	(43.66)	(58.19)	(134.88)
Within school variance	3 248.2***	4 120.5***	3 424.4***	3 601.1***	2 698.1***	2 438.3***	3 814.9***	3 233.7***	3 563.3***
	(299.20)	(328.19)	(356.69)	(263.50)	(260.93)	(260.58)	(334.44)	(302.33)	(542.33)
Number of observations	5 908	5 282	19 204	9 073	4 602	5 622	7 038	5 803	33 806

Individual features	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Native speaker	30.93**	-3.408	24.64***	10.23	4.315	-6.746*	23.48***	6.607	16.41*
	(3.12)	(-0.89)	(5.81)	(1.10)	(0.74)	(-2.06)	(3.50)	(1.42)	(2.23)
Constant	379.4***	422.5***	346.5***	434.6***	445.1***	430.1***	365.1***	442.1***	495.6***
	(30.53)	(94.45)	(62.18)	(44.62)	(57.03)	(106.76)	(47.25)	(59.42)	(58.46)
Between school variance	2 463.0***	2 081.8***	3 046.2***	2 979.6***	3 771.8***	2 927.4***	3 003.3***	5 068.5***	3 797.1***
	(26.01)	(55.60)	(68.77)	(72.05)	(64.62)	(64.01)	(49.34)	(66.90)	(65.73)
Within school variance	4 307.3***	4 450.8***	3 836.4***	3 587.2***	4 429.6***	3 932.3***	3 091.6***	3 171.5***	3 503.9***
	(205.44)	(260.24)	(317.05)	(299.25)	(270.05)	(300.48)	(299.17)	(269.15)	(260.38)
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use randon slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295644

[Part 1/1]

Table B.4 **Model 4: Rural-urban disparities**

School location	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Village (ref. town)	-22.09	-61.23***	-43.65***	-32.14***	-21.77**	-11.80	-2.128	-9.724	-32.21***
	(-1.42)	(-3.76)	(-5.55)	(-3.43)	(-2.87)	(-1.08)	(-0.28)	(-1.06)	(-7.69)
City	14.39	44.14***	27.70***	17.74*	13.44	36.05**	35.56***	15.45	25.36***
	(1.84)	(4.03)	(5.13)	(2.39)	(1.04)	(3.28)	(4.63)	(1.69)	(7.24)
Constant	385.3***	425.6***	377.9***	371.1***	410.4***	371.2***	370.7***	427.7***	407.5***
	(70.52)	(71.90)	(108.26)	(60.85)	(85.14)	(64.05)	(80.24)	(55.84)	(165.75)
Bewteen school variance	2 490.9***	4 003.1***	2 413.7***	1 687.1***	1 857.5***	2 374.6***	1 849.4***	1 728.4***	1 570.2***
	(84.36)	(87.16)	(77.60)	(59.30)	(48.68)	(41.41)	(46.81)	(54.82)	(138.88)
Within school variance	3 253.1***	4 160.1***	3 433.6***	3 617.4***	2 699.3***	2 456.3***	3 852.1***	3 235.4***	3 577.2***
	(300.07)	(321.15)	(357.84)	(266.68)	(261.22)	(255.97)	(325.30)	(301.57)	(540.12)
Number of observations	5 908	5 282	19 204	9 073	4 602	5 622	7 038	5 803	33 806

School location	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Village (ref. town)	m	-36.31***	-45.77***	-31.41**	-134.7***	-5.717	-11.76	27.84	-25.10*
		(-5.50)	(-5.45)	(-2.64)	(-10.06)	(-0.51)	(-0.95)	(0.87)	(-2.46)
City	13.45	25.75**	43.39***	27.95**	32.43**	33.92***	26.38*	-16.21	36.44*
	(0.52)	(2.69)	(5.32)	(2.86)	(2.96)	(3.65)	(1.99)	(-1.28)	(2.49)
Constant	405.7***	418.5***	359.3***	437.5***	437.3***	416.4***	382.3***	456.8***	514.0***
	(50.03)	(93.97)	(62.99)	(87.68)	(74.83)	(96.06)	(79.17)	(43.72)	(61.57)
Bewteen school variance	2 447.3***	1 778.2***	2 185.4***	2 693.5***	3 461.2***	2 572.9***	2 877.2***	5 009.6***	3 211.9***
	(29.17)	(56.30)	(60.56)	(79.19)	(62.88)	(64.14)	(53.02)	(66.05)	(63.85)
Within school variance	4 323.6***	4 449.3***	3 864.2***	3 588.4***	4 430.3***	3 942.0***	3 104.3***	3 172.8***	3 507.8***
	(206.31)	(259.85)	(310.30)	(299.41)	(269.54)	(303.08)	(302.84)	(268.97)	(260.30)
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use randon slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295650

[Part 1/1]

Table B.5 **Model 5: Mathematics performance and students' origin**

Individual features	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	12.30***	11.80***	9.267***	14.91***	13.44***	9.451***	12.21***	15.68***	6.930***
	(9.03)	(8.94)	(6.04)	(6.45)	(8.31)	(3.82)	(7.67)	(7.67)	(7.92)
ESCS squared	1.643*	-0.494	0.169	1.349	1.401*	0.757	0.136	-0.0 613	0.730*
	(2.29)	(-0.65)	(0.34)	(1.94)	(2.34)	(1.23)	(0.20)	(-0.04)	(2.19)
Gender (ref. male)	-15.36***	-13.33***	-19.73***	-23.95***	-23.32***	-5.527**	0.726	-2.302	-17.34***
	(-6.49)	(-6.22)	(-13.37)	(-11.06)	(-13.34)	(-3.05)	(0.10)	(-1.30)	(-17.50)
Native speaker	8.884	18.85***	12.48**	28.54***	6.186	-12.83***	18.15***	2.770	17.98***
	(1.76)	(4.07)	(2.66)	(4.33)	(0.93)	(-5.58)	(4.22)	(0.60)	(6.00)
Village (ref. town)	-15.01	-39.67*	-37.97***	-23.12**	-15.17*	-10.72	1.472	-6.052	-27.30***
	(-1.08)	(-2.55)	(-4.94)	(-2.64)	(-2.29)	(-1.03)	(0.20)	(-0.68)	(-6.47)
City	13.39	36.24***	21.38***	12.51	9.435	34.99**	31.43***	11.72	20.68***
	(1.92)	(3.61)	(4.29)	(1.89)	(0.82)	(3.29)	(4.52)	(1.35)	(6.18)
Constant	391.2***	421.8***	390.0***	372.7***	425.7***	392.7***	361.3***	431.8***	406.4***
	(56.78)	(59.77)	(61.49)	(37.90)	(56.12)	(63.11)	(47.52)	(52.32)	(113.53)
Between school variance	1 958.5***	3 252.0***	2 040.2***	1 264.9***	1 392.2***	2 188.2***	1 498.2***	1 553.8***	1 403.7***
	(76.61)	(82.28)	(74.66)	(56.56)	(45.37)	(42.41)	(39.85)	(48.69)	(133.37)
Within school variance	3 111.2***	3 991.8***	3 262.1***	3 358.8***	2 464.9***	2 390.7***	3 647.6***	3 129.9***	3 459.9***
	(308.64)	(324.71)	(358.39)	(261.62)	(264.62)	(258.58)	(342.40)	(302.80)	(534.24)
% of the between school varaince explained	24.60	30.02	24.90	35.19	29.82	17.90	30.80	16.48	27.59
% of the within school variance explained	4.36	4.05	5.01	7.16	8.69	2.67	5.31	3.23	3.31
Number of observations	5 819	5 183	18 951	8 997	4 571	5 606	6 908	5 797	33 598

Individual features	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS*	11.96***	21.16***	12.41***	19.32***	8.518***	17.21***	12.41***	3.108	10.22***
	(5.61)	(9.68)	(7.06)	(10.49)	(5.94)	(8.88)	(7.32)	(1.67)	(4.20)
ESCS squared	-0.0722	2.781**	0.752	1.780*	-0.651	3.087***	2.369***	-1.246	1.038
	(-0.06)	(3.27)	(1.35)	(2.23)	(-0.53)	(4.30)	(4.16)	(-1.79)	(1.47)
Gender (ref. male)	-10.75**	3.784	-26.06***	-12.13***	-24.09***	0.0 628	-23.61***	-21.67***	-23.91***
	(-2.99)	(1.80)	(-13.09)	(-5.02)	(-9.43)	(0.03)	(-12.38)	(-11.30)	(-11.21)
Native speaker	14.91	-5.096	24.83***	6.240	0.788	-11.48***	10.98	2.672	14.53
	(0.93)	(-1.44)	(5.99)	(0.73)	(0.13)	(-3.59)	(1.35)	(0.56)	(1.95)
Village (ref. town)	m	-28.00***	-37.57***	-24.08*	-128.0***	-3.530	-9.143	30.57	-22.47*
		(-4.69)	(-5.17)	(-2.03)	(-9.67)	(-0.31)	(-0.75)	(1.04)	(-2.20)
City	11.50	18.56*	33.80***	19.62*	29.38**	28.79***	21.98	-17.03	29.19*
	(0.48)	(2.24)	(4.69)	(2.31)	(2.75)	(3.48)	(1.75)	(-1.38)	(2.01)
Constant	400.6***	431.6***	365.7***	447.0***	453.1***	437.2***	393.3***	474.3***	527.0***
	(25.34)	(87.33)	(58.93)	(46.21)	(49.14)	(91.20)	(40.10)	(41.76)	(44.40)
Between school variance	2 066.2***	1 270.0***	1 645.6***	2 076.2***	3 311.5***	2 200.3***	2 631.6***	4 738.9***	3 134.0***
	(29.04)	(49.91)	(53.67)	(72.62)	(62.02)	(58.88)	(52.93)	(64.54)	(63.79)
Within school variance	4 188.8***	4 273.5***	3 610.0***	3 387.1***	4 250.4***	3 845.6***	2 875.2***	3 031.9***	3 308.6***
	(207.01)	(281.43)	(307.29)	(304.55)	(273.40)	(312.93)	(287.74)	(260.39)	(261.28)
% of the between school varaince explained	16.86	40.34	49.26	30.48	12.29	23.21	12.77	7.23	18.04
% of the within school variance explained	3.12	3.95	6.59	5.62	4.07	2.42	7.38	4.44	5.67
Number of observations	4 688	5 183	6 005	5 055	4 624	6 586	4 339	4 806	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295661

[Part 1/1]

Table B.6 **Model 6: Socio-economic background of the students and schools' socio-economic composition**

Individual features	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	11.29***	11.51***	8.232***	13.69***	12.55***	7.659**	10.87***	14.76***	5.873***
	(8.16)	(8.56)	(5.53)	(5.79)	(7.69)	(2.99)	(6.57)	(7.11)	(6.78)
ESCS squared	1.555*	-0.110	0.0 223	1.001	1.085	0.614	-0.0 452	-0.208	0.380
	(2.19)	(-0.14)	(0.05)	(1.47)	(1.78)	(0.96)	(-0.06)	(-0.15)	(1.18)
School's mean ESCS	49.03***	72.91***	45.70***	35.10***	33.49***	36.97***	47.13***	45.47***	29.54***
	(8.91)	(13.00)	(16.14)	(9.86)	(9.42)	(6.11)	(5.48)	(5.68)	(20.40)
Constant	430.0***	462.5***	455.2***	435.4***	449.9***	452.8***	411.6***	450.3***	452.2***
	(114.26)	(148.02)	(121.68)	(95.90)	(101.86)	(39.99)	(83.27)	(121.57)	(239.01)
Between school variance	983.5***	1 294.3***	1 022.4***	762.7***	749.9***	1 785.9***	1 239.6***	1 307.8***	1 043.0***
	(49.51)	(46.92)	(75.00)	(49.98)	(48.82)	(49.29)	(41.99)	(37.14)	(118.74)
Within school variance	3 170.5***	4 057.7***	3 366.8***	3 507.4***	2 603.1***	2 423.6***	3 666.9***	3 133.2***	3 545.4***
	(308.62)	(320.14)	(351.56)	(270.09)	(268.92)	(252.13)	(338.58)	(301.66)	(542.55)
% of the between school variance explained	62.14	72.15	62.36	60.92	62.20	32.99	42.74	29.70	46.20
% of the within school variance explained	2.54	2.47	1.97	3.05	3.57	1.33	4.81	3.13	0.92
Number of observations	5 819	5 183	18 951	8 997	4 571	5 606	6 908	5 797	33 598

Individual features	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS	12.27***	19.46***	10.57***	19.00***	8.751***	15.66***	12.67***	3.097	10.79***
	(6.28)	(8.91)	(6.27)	(10.17)	(5.79)	(7.92)	(6.97)	(1.62)	(4.34)
ESCS squared	-0.0 104	2.704**	0.297	2.186*	-0.692	3.102***	2.540***	-1.132	0.988
	(-0.01)	(3.21)	(0.55)	(2.38)	(-0.54)	(4.30)	(4.25)	(-1.62)	(1.36)
School's mean ESCS	102.1***	48.40***	49.03***	56.86***	101.7***	34.64***	44.33***	83.27***	49.23***
	(16.72)	(8.33)	(18.52)	(9.74)	(14.46)	(7.06)	(8.13)	(11.22)	(7.59)
Constant	439.1***	465.9***	441.0***	477.8***	482.9***	483.9***	449.2***	577.8***	615.8***
	(126.64)	(85.24)	(128.85)	(128.45)	(117.21)	(70.20)	(58.95)	(45.33)	(49.89)
Between school variance	359.2***	875.0***	697.0***	1 109.0***	1 291.8***	1 651.3***	1 513.0***	2 197.4***	2 033.2***
	(22.10)	(53.79)	(49.23)	(46.62)	(50.15)	(40.81)	(59.71)	(61.83)	(62.33)
Within school variance	4 219.8***	4 273.4***	3 792.4***	3 425.7***	4 385.5***	3 867.4***	3 021.2***	3 126.5***	3 458.8***
	(218.34)	(277.87)	(305.58)	(302.51)	(273.72)	(314.44)	(295.73)	(272.36)	(264.44)
% of the between school variance explained	85.55	58.90	78.51	62.87	65.79	42.37	49.85	56.98	46.82
% of the within school variance explained	2.40	3.95	1.87	4.54	1.02	1.87	2.67	1.46	1.39
Number of observations	4 688	5 183	6 005	5 055	4 624	6 586	4 339	4 806	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295674

[Part 1/1]

Table B.7 **Model 7: Performance diffrences associated with national study programme**

	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Vocational track (ref. general)	17.77	-38.70***	-8.128	45.05***	35.06***	17.08	m	-17.15*	22.22***
	(1.84)	(-4.02)	(-0.67)	(10.11)	(9.88)	(1.70)		(-2.02)	(7.03)
Constant	386.2***	455.2***	389.3***	365.5***	404.0***	371.7***	385.7***	432.7***	408.2***
	(88.63)	(64.10)	(144.72)	(101.02)	(104.10)	(70.82)	(103.06)	(120.10)	(204.27)
Between school variance	2 596.8***	4 270.0***	2 716.5***	2 042.4***	2 025.3***	2 618.3***	2 165.0***	1 838.7***	1 847.1***
	(84.32)	(82.17)	(76.68)	(56.94)	(48.18)	(46.35)	(43.89)	(58.49)	(131.40)
Within school variance	3 250.7***	4 160.8***	3 434.2***	3 395.0***	2 646.0***	2 456.3***	3 852.1***	3 234.8***	3 578.0***
	(297.78)	(320.99)	(357.93)	(273.33)	(263.36)	(255.97)	(325.30)	(301.68)	(540.06)
% of the between school variance explained	0.03	8.11	0.00	-4.64	-2.10	1.76	0.00	1.16	4.71
% of the within school variance explained	0.08	-0.01	0.00	6.16	1.98	0.00	0.00	-0.01	0.01
Number of observations	5 908	5 282	19 204	9 073	4 602	5 622	7 038	5 803	33 806

	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Vocational track (ref. general)	-68.09***	-12.10*	m	m	-80.97***	-62.82***	m	-63.08***	-79.54***
	(-9.40)	(-2.11)			(-9.64)	(-7.97)		(-6.83)	(-14.50)
Constant	454.9***	422.2***	368.7***	444.6***	509.8***	438.9***	388.0***	472.4***	512.2***
	(47.38)	(105.03)	(88.89)	(101.92)	(63.11)	(94.96)	(83.34)	(56.89)	(99.35)
Between school variance	1 269.4***	2 109.1***	3 243.4***	2 986.6***	2 187.4***	2 393.2***	3 017.0***	4 152.6***	3 780.1***
	(27.46)	(57.89)	(71.55)	(72.13)	(55.18)	(59.35)	(49.51)	(69.31)	(65.12)
Within school variance	4 158.7***	4 442.3***	3 864.6***	3 588.7***	4 395.1***	3 924.8***	3 104.2***	3 173.0***	3 507.6***
	(264.49)	(259.40)	(310.45)	(299.45)	(267.18)	(312.36)	(302.83)	(268.94)	(260.27)
% of the between school variance explained	48.92	0.93	0.00	0.00	42.07	16.47	0.00	18.70	1.14
% of the within school variance explained	3.81	0.16	0.00	0.00	0.80	0.41	0.00	-0.01	0.00
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295684

[Part 1/1]

Table B.8 **Model 8: Performance of private schools (government-dependent and government-independent schools combined) vs. public schools**

Type of school	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Private (ref. public)	57.85***	108.2*	80.15***	49.81***	77.99***	-4.472	60.13***	2.541	41.68***
	(8.51)	(2.24)	(9.39)	(5.08)	(7.73)	(-0.47)	(4.67)	(0.22)	(7.55)
Constant	370.4***	438.1***	379.6***	370.2***	397.4***	377.0***	375.7***	431.4***	410.2***
	(88.15)	(84.82)	(156.55)	(118.19)	(117.28)	(66.73)	(128.23)	(123.72)	(239.82)
Between school variance	1 862.7***	4 505.6***	2 010.3***	1 645.0***	1 293.6***	2 660.2***	1 660.4***	1 859.9***	1 798.7***
	(68.07)	(84.71)	(69.48)	(61.74)	(45.78)	(47.01)	(51.66)	(58.16)	(137.75)
Within school variance	3 253.5***	4 160.3***	3 434.7***	3 618.7***	2 700.4***	2 456.3***	3 852.3***	3 234.8***	3 578.2***
	(300.02)	(321.05)	(357.97)	(266.61)	(260.88)	(255.97)	(325.29)	(301.66)	(540.08)
% of the between school variance explained	32.81	2.84	41.74	31.38	35.08	10.63	26.70	0.04	9.14
% of the within school variance explained	0.03	0.50	0.84	-0.12	0.06	-0.43	-0.01	-0.01	0.02
Number of observations	5 908	5 247	19 204	9 073	4 602	5 622	7 038	5 803	33 806

Type of school	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Private (ref. public)	-38.99***	87.29***	65.91***	70.25***	4.990	-35.61***	-58.27***	m	-13.56
	(-3.47)	(3.95)	(5.54)	(13.40)	(0.21)	(-3.50)	(-5.79)		(-0.85)
Constant	409.9***	417.6***	360.2***	444.2***	449.1***	432.4***	388.3***	448.2***	512.7***
	(42.46)	(113.29)	(91.16)	(101.71)	(85.05)	(97.87)	(83.04)	(74.14)	(94.12)
Between school variance	2 480.1***	1 880.5***	2 735.1***	2 954.3***	3 775.5***	2 691.6***	3 001.6***	5 108.0***	3 810.0***
	(26.04)	(59.56)	(73.51)	(70.80)	(64.59)	(61.65)	(49.40)	(67.16)	(65.81)
Within school variance	4 323.6***	4 449.3***	3 865.3***	3 588.8***	4 430.6***	3 941.1***	3 104.1***	3 172.7***	3 507.6***
	(206.31)	(259.85)	(310.42)	(299.44)	(269.47)	(303.15)	(302.85)	(268.99)	(260.27)
% of the between school variance explained	0.21	11.04	33.33	1.08	4.35	6.07	2.05	0.00	0.29
% of the within school variance explained	0.00	0.97	2.13	0.00	0.15	0.00	-0.07	0.75	-0.54
Number of observations	4 744	5 197	6 035	5 074	4 684	6 606	4 407	4 848	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295696

[Part 1/1]

Table B.9 **Model 8a: Performance in private schools (government-dependent and government-independent schools combined) vs. public schools, controlled for socio-economic background**

Type of school	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
Private (ref. public)	25.86***	-9.310	12.02	6.303	15.32	-3.584	32.75***	-6.127	-15.38***
	(3.65)	(-0.25)	(1.76)	(0.91)	(1.41)	(-0.50)	(3.46)	(-0.63)	(-3.55)
ESCS*	11.24***	11.52***	8.142***	13.63***	12.38***	7.661**	10.78***	14.76***	6.004***
	(8.19)	(8.56)	(5.47)	(5.74)	(7.61)	(3.00)	(6.51)	(7.11)	(6.90)
ESCS squared	1.521*	-0.0 982	-0.0 158	0.977	1.009	0.615	-0.116	-0.208	0.450
	(2.16)	(-0.13)	(-0.03)	(1.43)	(1.65)	(0.96)	(-0.17)	(-0.15)	(1.39)
Mean escs of the school	41.25***	73.27***	42.67***	33.79***	29.26***	36.94***	36.68***	45.63***	32.10***
	(7.30)	(12.82)	(12.62)	(8.82)	(6.31)	(6.00)	(4.54)	(5.70)	(20.11)
Constant	416.3***	462.7***	450.0***	432.9***	443.8***	454.2***	401.8***	450.5***	456.3***
	(74.43)	(147.51)	(92.10)	(79.39)	(72.73)	(40.22)	(82.39)	(119.63)	(205.21)
Between school variance	858.9***	1 292.4***	1 012.9***	757.8***	732.6***	1 782.5***	1 115.8***	1 307.1***	1 027.7***
	(53.60)	(46.77)	(73.02)	(50.42)	(45.82)	(49.49)	(44.85)	(37.19)	(116.58)
Within school variance	3 171.1***	4 057.8***	3 366.7***	3 507.5***	2 603.4***	2 423.6***	3 667.1***	3 133.2***	3 545.5***
	(308.52)	(320.14)	(351.67)	(270.09)	(268.78)	(252.14)	(338.60)	(301.66)	(542.50)
% of the between school variance explained	66.96	72.19	62.74	61.20	63.09	33.13	48.50	29.73	47.01
% of the within school variance explained	2.52	2.48	1.98	3.04	3.56	1.34	4.80	3.12	0.92
Number of observations	5 819	5 183	18 951	8 997	4 571	5 606	6 908	5 797	33 598

Type of school	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
Private (ref. public)	-85.24***	38.20**	4.601	-39.39***	-81.34***	-40.24***	-94.78**	0	-58.05***
	(-12.23)	(2.72)	(0.68)	(-4.41)	(-7.75)	(-6.01)	(-2.65)	c	(-3.52)
ESCS	12.27***	19.42***	10.55***	19.03***	8.760***	15.72***	12.69***	3.097	10.75***
	(6.28)	(8.87)	(6.27)	(10.17)	(5.79)	(7.98)	(6.98)	(1.62)	(4.31)
ESCS squared	-0.00779	2.682**	0.287	2.212*	-0.671	3.131***	2.548***	-1.132	0.972
	(-0.01)	(3.18)	(0.54)	(2.39)	(-0.53)	(4.37)	(4.26)	(-1.62)	(1.33)
Mean escs of the school	102.7***	45.12***	48.25***	57.96***	102.7***	36.04***	44.86***	83.27***	54.42***
	(17.18)	(7.93)	(18.12)	(9.72)	(14.63)	(7.65)	(8.34)	(11.22)	(10.60)
Constant	439.5***	462.2***	439.5***	478.6***	483.5***	492.4***	450.3***	577.8***	629.5***
	(128.48)	(84.90)	(126.16)	(127.32)	(117.41)	(75.29)	(59.99)	(45.33)	(69.17)
Between school variance	344.7***	830.3***	694.7***	1 097.9***	1 273.5***	1 416.9***	1 481.0***	2 197.4***	1 808.1***
	(21.78)	(52.81)	(49.71)	(45.61)	(49.70)	(37.36)	(60.18)	(61.83)	(54.31)
Within school variance	4 219.4***	4 273.4***	3 792.4***	3 425.9***	4 385.2***	3 868.2***	3 020.5***	3 126.5***	3 459.2***
	(218.35)	(277.86)	(305.60)	(302.49)	(273.78)	(314.58)	(295.81)	(272.36)	(264.49)
% of the between school variance explained	86.16	61.00	78.60	63.26	66.28	50.58	50.91	56.99	52.71
% of the within school variance explained	2.41	3.96	1.86	4.54	1.02	1.85	2.71	1.45	1.37
Number of observations	4 688	5 183	6 005	5 055	4 624	6 586	4 339	4 806	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
c: There are too few observations or no observation to provide reliable estimates (i.e. there are fewer than 30 students or fewer than 5 schools with valid data).
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295703

[Part 1/1]

Table B.10 Model 9: Relation between individual features, type of school, national study programme and performance

	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	10.88***	10.57***	7.449***	12.35***	11.48***	8.527***	11.23***	14.88***	5.842***
	(8.05)	(7.95)	(5.13)	(5.19)	(7.45)	(3.31)	(6.74)	(7.12)	(6.66)
ESCS squared	1.626*	0.0873	-0.0231	1.380*	1.260*	0.656	-0.00191	-0.129	0.763*
	(2.33)	(0.12)	(-0.05)	(2.03)	(2.17)	(1.03)	(-0.00)	(-0.09)	(2.29)
Gender (ref. male)	-15.32***	-13.72***	-19.70***	-25.98***	-23.83***	-5.604**	8.101	-2.376	-17.43***
	(-6.45)	(-6.34)	(-13.32)	(-13.12)	(-13.61)	(-3.10)	(1.22)	(-1.35)	(-17.61)
Native speaker	8.845	15.76***	12.31**	24.60***	5.913	-13.20***	18.68***	2.838	17.14***
	(1.76)	(3.40)	(2.60)	(3.53)	(0.89)	(-5.74)	(4.36)	(0.62)	(5.76)
Village (ref. town)	7.797	-13.91	5.190	11.23	0.946	15.83	9.875	17.04*	2.956
	(1.15)	(-1.75)	(1.01)	(1.45)	(0.19)	(1.56)	(1.31)	(1.97)	(0.74)
City	7.418	-3.987	-11.92**	-2.877	-7.096	16.81	17.32**	13.41	-3.185
	(1.57)	(-0.59)	(-3.12)	(-0.43)	(-0.74)	(1.47)	(2.62)	(1.55)	(-0.93)
Vocational track (ref. general)	20.87**	-0.852	-10.70	42.60***	35.49***	26.44*	m	-13.45	19.86***
	(3.11)	(-0.13)	(-0.82)	(10.75)	(10.94)	(2.39)		(-1.76)	(7.72)
School type: private (ref. public)	29.01***	-4.366	9.475	11.53	16.30	-5.722	28.94***	7.550	-10.78*
	(1.57)	(-0.59)	(-3.12)	(-0.43)	(-0.74)	(1.47)	(2.62)	(1.55)	(-0.93)
Constant	407.5***	458.9***	460.1***	413.1***	449.2***	453.8***	370.4***	436.5***	444.1***
	(51.39)	(67.13)	(56.92)	(35.36)	(48.29)	(36.33)	(39.49)	(47.18)	(93.54)
Between school variance	789.5***	1 314.5***	1 003.6***	800.8***	727.6***	1 598.0***	958.7***	1 278.3***	960.2***
	(53.38)	(47.64)	(72.82)	(50.43)	(42.08)	(54.35)	(37.34)	(37.32)	(116.01)
Within school variance	3 108.7***	3 997.1***	3 263.0***	3 143.7***	2 410.7***	2 390.5***	3 654.6***	3 130.1***	3 460.3***
	(305.89)	(322.54)	(359.34)	(261.33)	(264.73)	(258.80)	(343.38)	(302.51)	(533.06)
% of the between school varaince explained	69.62	71.72	63.07	59.00	63.34	40.04	55.75	31.29	50.46
% of the within school variance explained	4.46	3.92	4.98	13.10	10.71	2.69	5.14	3.22	3.30
Number of observations	5 819	5 183	18 951	8 997	4 571	5 606	6 908	5 797	33 598

	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS	9.362***	19.47***	10.83***	18.22***	6.676***	16.32***	11.61***	2.331	9.730***
	(5.27)	(8.91)	(6.45)	(9.73)	(4.75)	(8.30)	(6.70)	(1.25)	(3.97)
ESCS squared	-0.640	2.679**	0.844	2.122*	-0.957	3.113***	2.362***	-1.268	1.117
	(-0.54)	(3.18)	(1.58)	(2.32)	(-0.78)	(4.34)	(4.15)	(-1.82)	(1.57)
Gender (ref. male)	-12.47***	3.600	-25.98***	-12.43***	-24.16***	-0.729	-23.65***	-21.70***	-24.01***
	(-3.78)	(1.72)	(-13.08)	(-5.15)	(-9.52)	(-0.36)	(-12.38)	(-11.33)	(-11.26)
Native speaker	15.74	-4.009	24.25***	7.168	-0.143	-12.11***	11.07	1.686	12.82
	(0.95)	(-1.14)	(6.02)	(0.84)	(-0.02)	(-3.83)	(1.36)	(0.35)	(1.71)
Village (ref. town)	-14.95	-7.437	0.940	8.957	-11.97	9.022	3.380	33.41*	11.39
	(-1.85)	(-1.26)	(0.14)	(1.21)	(-1.33)	(1.34)	(0.40)	(2.30)	(1.20)
City	11.12	-3.116	-1.979	-2.269	1.325	8.797	-5.158	-15.64*	-17.51
	(1.23)	(-0.44)	(-0.33)	(-0.36)	(0.19)	(1.39)	(-0.60)	(-2.00)	(-1.41)
Vocational track (ref. general)	-57.36***	-9.312	m	m	-47.14***	-38.79***	m	-27.72***	-65.59***
	(-7.14)	(-1.79)			(-5.26)	(-4.40)		(-4.23)	(-12.69)
School type: private (ref. public)	-90.26***	35.99*	1.645	-42.29***	-95.59***	-24.21***	-101.3**	m	-57.51**
	(1.23)	(-0.44)	(-0.33)	(-0.36)	(0.19)	(1.39)	(-0.60)		(-1.41)
Constant	456.9***	465.9***	432.2***	478.1***	523.6***	497.6***	455.1***	600.1***	648.6***
	(31.66)	(70.88)	(55.77)	(49.54)	(53.19)	(56.01)	(39.89)	(44.63)	(40.27)
Between school variance	353.5***	790.4***	683.7***	1 094.0***	1 274.2***	1 238.4***	1 549.4***	1 771.2***	1 835.5***
	(16.03)	(53.25)	(50.81)	(47.10)	(50.00)	(34.09)	(59.62)	(59.19)	(50.03)
Within school variance	4 045.4***	4 266.7***	3 612.5***	3 388.0***	4 218.3***	3 837.8***	2 874.7***	3 031.6***	3 309.0***
	(252.23)	(278.52)	(306.55)	(303.27)	(273.06)	(319.77)	(287.83)	(260.29)	(260.78)
% of the between school varaince explained	85.79	62.88	78.94	63.36	66.25	56.79	48.66	65.33	52.00
% of the within school variance explained	6.43	4.11	6.52	5.57	4.79	2.64	7.41	4.45	5.65
Number of observations	4 688	5 183	6 005	5 055	4 624	6 586	4 339	4 806	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295715

[Part 1/2]

Table B.11 **Model 10: Grade repetition, preschool attendance and mathematics performance**

	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	9.812***	11.08***	7.467***	9.968***	10.62***	8.805***	10.14***	14.44***	6.795***
	(7.74)	(8.52)	(5.02)	(4.48)	(7.24)	(3.45)	(6.95)	(7.14)	(7.96)
ESCS squared	1.507*	-0.0 704	0.443	1.088	1.271*	0.767	0.450	0.120	0.898**
	(2.23)	(-0.10)	(0.92)	(1.75)	(2.33)	(1.22)	(0.69)	(0.09)	(2.79)
Gender (ref. male)	-20.61***	-15.48***	-26.53***	-30.38***	-28.03***	-7.391***	-2.529	-3.707*	-19.25***
	(-9.41)	(-7.46)	(-18.39)	(-16.89)	(-16.63)	(-4.17)	(-0.33)	(-2.08)	(-20.44)
Native speaker	4.348	17.65***	8.551*	18.94**	4.566	-12.19***	16.39***	2.632	17.17***
	(1.01)	(4.17)	(2.09)	(2.78)	(0.76)	(-5.24)	(4.20)	(0.57)	(5.90)
Grade 7	-83.51***	m	-74.89***	-69.16***	m	-57.39***	-53.22	m	-67.50***
	(-10.50)		(-20.82)	(-14.91)		(-4.64)	(-1.32)		(-10.60)
Grade 8	-58.14***	-69.24***	-56.22***	-46.91***	-45.03***	-50.73***	-23.44	-20.99	-47.54***
	(-14.28)	(-3.87)	(-16.39)	(-13.46)	(-13.15)	(-4.52)	(-1.70)	(-0.93)	(-10.79)
Grade 9	-30.04***	-45.01***	-29.19***	-30.08***	-21.45***	-32.93***	-25.17***	-18.53***	-25.54***
	(-8.40)	(-5.69)	(-15.66)	(-9.38)	(-7.27)	(-3.29)	(-4.19)	(-4.04)	(-7.31)
Grade 11	9.941	20.04***	26.69***	29.26***	19.18***	9.203*	m	18.21***	24.34***
	(1.51)	(3.51)	(7.97)	(11.20)	(8.37)	(2.09)		(7.23)	(4.87)
Grade 12	18.06	-3.250	m	m	63.48***	43.47*	m	38.42	25.95
	(1.33)	(-0.08)			(3.65)	(1.97)		(1.85)	(0.85)
Repeated a grade	-12.26***	-33.63***	-13.99***	-3.267	-13.02***	-17.48***	-35.91***	-15.40*	-32.43***
	(-3.64)	(-4.66)	(-7.84)	(-1.52)	(-5.45)	(-6.89)	(-7.27)	(-2.00)	(-15.04)
Vocational track (ref. general)	13.87	-39.81***	-28.89*	1.115	1.951	-7.635	m	-31.78***	1.140
	(1.71)	(-4.76)	(-2.13)	(0.28)	(0.48)	(-0.63)		(-4.11)	(0.37)
Attended preschool	28.00***	12.32**	9.199***	4.091	8.431**	3.455	18.27***	0.972	15.56***
	(7.63)	(3.19)	(6.25)	(1.41)	(3.12)	(1.39)	(7.26)	(0.41)	(8.60)
Constant	391.6***	443.2***	423.8***	391.1***	420.8***	415.6***	368.2***	433.6***	413.5***
	(64.70)	(51.94)	(77.04)	(51.43)	(67.58)	(44.26)	(46.34)	(78.75)	(109.10)
Between school variance	1 339.2***	2 974.5***	1 502.4***	1 194.4***	1 050.4***	1 977.1***	1 625.7***	1 676.9***	1 084.0***
	(64.68)	(76.99)	(61.42)	(53.45)	(38.95)	(50.95)	(37.09)	(52.92)	(107.21)
Within school variance	2 662.7***	3 871.3***	2 715.4***	2 654.7***	2 073.9***	2 322.2***	3 399.4***	3 050.4***	3 310.2***
	(267.22)	(326.59)	(381.65)	(259.99)	(262.06)	(243.21)	(364.57)	(304.57)	(583.88)
% of the between school variance explained	48.44	35.99	44.70	38.81	47.05	25.82	24.91	9.86	44.08
% of the within school variance explained	18.15	6.95	20.93	26.62	23.17	5.46	11.75	5.69	7.49
Number of observations	5 819	5 183	18 951	8 997	4 571	5 606	6 908	5 797	33 598

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295721

[Part 2/2]

Table B.11 **Model 10: Grade repetition, preschool attendance and mathematics performance**

	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS*	9.252***	19.40***	10.08***	19.21***	7.114***	16.31***	10.84***	1.741	10.81***
	(5.42)	(9.14)	(5.78)	(10.52)	(5.18)	(8.57)	(6.38)	(0.94)	(4.49)
ESCS squared	-0.568	2.697**	1.088*	1.806*	-1.055	2.862***	2.207***	-1.353*	1.222
	(-0.48)	(3.20)	(2.06)	(2.29)	(-0.86)	(4.04)	(3.98)	(-1.99)	(1.72)
Gender (ref. male)	-13.34***	0.925	-29.39***	-12.29***	-24.42***	-1.744	-26.45***	-27.12***	-24.69***
	(-4.02)	(0.44)	(-15.26)	(-5.12)	(-9.62)	(-0.87)	(-15.06)	(-15.27)	(-11.99)
Native speaker	11.04	-11.39***	20.03***	5.248	0.152	-9.928**	11.62	1.744	14.27*
	(0.69)	(-3.37)	(5.68)	(0.61)	(0.02)	(-3.14)	(1.46)	(0.40)	(2.05)
Grade 7	m	m	-69.02***	m	m	-0.0 588	-106.5***	-103.9***	-151.9***
			(-9.66)			(-0.00)	(-15.65)	(-5.14)	(-9.99)
Grade 8	m	-154.1***	-46.90***	-0.765	-163.8***	-22.92	-70.21***	-105.2***	-107.5***
		(-3.58)	(-12.09)	(-0.03)	(-3.74)	(-1.59)	(-11.12)	(-7.15)	(-8.60)
Grade 9	-6.112	-88.04***	-24.74***	-1.577	-154.5***	-22.39***	-37.21***	-34.29***	-83.26***
	(-0.17)	(-13.39)	(-8.77)	(-0.10)	(-8.58)	(-6.86)	(-6.00)	(-11.12)	(-7.50)
Grade 11	11.82**	16.88	22.04***	10.03*	33.52***	20.51**	24.16***	20.50***	m
	(3.22)	(1.09)	(10.78)	(2.25)	(3.81)	(2.85)	(5.27)	(3.40)	
Grade 12	m	m	m	m	m	m	m	27.38	m
								(1.83)	
Repeated a grade	-33.74***	m	-12.05***	-17.92**	-40.50***	-20.93***	-22.70***	-23.52***	-13.60
	(-3.92)		(-5.02)	(-3.01)	(-3.33)	(-3.57)	(-5.05)	(-7.00)	(-1.85)
Vocational track (ref. general)	-64.54***	-14.72**	m	m	-82.98***	-63.09***	m	-65.78***	10.83
	(-8.61)	(-2.73)			(-9.31)	(-8.60)		(-7.66)	(0.88)
Attended preschool	0.260	11.29***	7.797**	14.76**	-1.479	30.44***	1.703	-0.00 499	7.860*
	(0.12)	(4.82)	(2.96)	(3.20)	(-0.53)	(3.74)	(0.79)	(-0.00)	(2.35)
Constant	449.5***	433.8***	376.2***	439.2***	529.8***	432.8***	425.0***	506.9***	532.7***
	(28.96)	(93.36)	(69.35)	(45.24)	(50.47)	(45.08)	(47.09)	(55.37)	(61.05)
Between school variance	1 109.2***	1 378.9***	1 843.9***	2 157.8***	1 837.9***	1 923.1***	1 037.7***	3 213.0***	1 915.4***
	(27.58)	(49.59)	(51.59)	(67.72)	(59.03)	(53.35)	(24.24)	(63.21)	(50.36)
Within school variance	4 009.4***	3 974.8***	3 093.9***	3 368.6***	4 184.5***	3 734.2***	2 563.6***	2 616.7***	3 260.9***
	(244.57)	(297.24)	(291.45)	(296.79)	(276.59)	(330.75)	(250.59)	(275.95)	(261.63)
% of the between school variance explained	55.37	35.23	43.15	27.75	51.32	32.88	65.60	37.10	49.91
% of the within school variance explained	7.27	10.66	19.94	6.13	5.55	5.25	17.42	17.52	7.03
Number of observations	4 688	5 183	6 005	5 055	4 624	6 586	4 339	4 806	4 959

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295721

[Part 1/2]

Table B.12 **Model 11: OTL and mathematics performance**

	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	9.356***	8.616***	6.429***	8.467**	10.61***	8.778**	10.83***	11.58***	6.407***
	(6.59)	(5.35)	(3.61)	(2.94)	(5.27)	(2.68)	(6.20)	(3.86)	(6.14)
ESCS squared	1.711*	0.447	0.468	1.169	1.332*	1.178	1.435*	0.487	0.927*
	(2.45)	(0.46)	(0.80)	(1.38)	(2.03)	(1.34)	(1.98)	(0.23)	(2.32)
Gender (ref. male)	-18.70***	-19.19***	-27.45***	-33.83***	-28.08***	-8.333***	-6.525	-6.627**	-20.24***
	(-7.02)	(-7.31)	(-16.87)	(-14.72)	(-12.01)	(-3.67)	(-0.92)	(-3.07)	(-16.77)
Native speaker	4.449	11.10*	10.87*	22.39**	8.865	-13.61***	17.47***	6.190	17.77***
	(0.72)	(2.24)	(2.34)	(3.08)	(1.27)	(-4.09)	(3.91)	(1.06)	(6.14)
Grade 7	-76.65***	m	-70.25***	-50.11***	m	-52.84***	-77.45*	m	-50.36***
	(-7.80)		(-14.32)	(-7.25)		(-4.28)	(-2.21)		(-5.33)
Grade 8	-53.12***	-42.05*	-51.64***	-35.38***	-43.98***	-48.83***	-4.793	-29.29	-36.83***
	(-9.79)	(-2.42)	(-11.15)	(-7.27)	(-7.49)	(-4.49)	(-0.32)	(-1.23)	(-7.40)
Grade 9	-27.27***	-34.51***	-26.03***	-21.38***	-23.38***	-32.99***	-17.60**	-9.724	-23.05***
	(-6.34)	(-3.37)	(-9.37)	(-4.55)	(-6.72)	(-3.42)	(-2.85)	(-1.79)	(-6.51)
Grade 11	5.509	9.775	25.74***	26.35***	11.96**	15.38*	m	12.64***	21.29***
	(0.75)	(1.49)	(6.64)	(8.00)	(2.62)	(2.50)		(4.06)	(3.77)
Grade 12	20.99	m	m	m	52.29***	43.26	m	46.88*	9.266
	(1.43)				(3.52)	(1.74)		(2.34)	(0.27)
Grade repetition	-9.931*	-34.25***	-11.74***	-2.285	-12.43***	-16.12***	-30.85***	-7.284	-27.12***
	(-2.08)	(-3.67)	(-4.91)	(-0.76)	(-3.45)	(-4.97)	(-5.43)	(-0.74)	(-11.65)
Vocational track (ref. general)	6.010	-38.32***	-32.50	4.489	6.284	-4.913	m	-28.11***	2.878
	(0.81)	(-4.86)	(-1.95)	(1.02)	(1.38)	(-0.42)		(-3.97)	(1.00)
Attended preschool	30.00***	11.74**	8.458***	5.467	7.786**	4.176	15.07***	1.417	16.30***
	(5.71)	(2.62)	(4.59)	(1.53)	(2.63)	(1.58)	(5.14)	(0.50)	(6.95)
Index of OTL: Formal mathematics	14.29***	27.90***	17.38***	19.30***	14.09***	10.43***	22.93***	18.74***	27.89***
	(7.91)	(13.18)	(10.95)	(8.28)	(6.34)	(4.63)	(12.80)	(7.99)	(24.33)
Constant	361.3***	316.6***	557.7***	534.3***	538.4***	533.5***	61.72	346.0***	501.5***
	(5.76)	(4.94)	(11.37)	(8.40)	(6.06)	(8.29)	(0.94)	(4.95)	(10.14)
Between school variance	1 159.1***	2 570.3***	1 227.5***	1 048.5***	910.9***	1 830.0***	1 403.5***	1 496.0***	864.0***
	(60.85)	(68.09)	(54.61)	(48.08)	(35.20)	(47.40)	(36.68)	(46.73)	(93.29)
Within school variance	2 576.8***	3 635.5***	2 671.0***	2 614.0***	2 076.8***	2 309.8***	3 138.8***	2 951.9***	3 095.4***
	(228.11)	(227.75)	(301.14)	(215.87)	(194.63)	(219.57)	(269.38)	(270.62)	(485.07)
% of the between school variance explained	55.38	44.69	54.81	46.28	54.08	31.34	35.17	19.58	55.43
% of the within school variance explained	20.79	12.61	22.23	27.74	23.06	5.96	18.52	8.74	13.49
Number of observations	3 779	3 360	12 217	5 663	2 886	3 692	4 540	3 844	22 209

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295733

[Part 2/2]

Table B.12 **Model 11: OTL and mathematics performance**

	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS*	6.862***	16.68***	11.30***	16.49***	5.507***	14.98***	10.92***	0.879	10.18**
	(3.68)	(6.49)	(5.79)	(7.28)	(3.42)	(7.12)	(5.18)	(0.32)	(3.18)
ESCS squared	-0.737	3.063**	1.488*	1.168	-1.881	3.383***	1.960**	-1.337	1.098
	(-0.55)	(3.01)	(2.24)	(1.29)	(-1.44)	(4.30)	(2.85)	(-1.38)	(1.18)
Gender (ref. male)	-18.12***	-2.373	-31.42***	-11.32***	-28.05***	-3.235	-26.41***	-29.64***	-23.32***
	(-5.55)	(-0.99)	(-13.18)	(-4.16)	(-8.80)	(-1.42)	(-10.89)	(-12.61)	(-9.43)
Native speaker	7.093	-10.93**	17.88***	4.537	1.303	-8.248*	14.19	5.998	10.74
	(0.46)	(-2.98)	(4.46)	(0.48)	(0.17)	(-2.27)	(1.53)	(0.99)	(1.59)
Grade 7	m	m	-54.68***	m	m	19.51	-112.7***	-89.73***	-131.4***
			(-6.49)			(0.50)	(-14.51)	(-3.57)	(-6.79)
Grade 8	m	-192.7***	-36.49***	-0.488	-150.5**	-10.24	-71.00***	-85.02***	-82.48***
		(-5.68)	(-7.14)	(-0.01)	(-2.81)	(-0.61)	(-8.75)	(-5.97)	(-5.40)
Grade 9	-11.11	-69.26***	-16.01***	1.315	-150.4***	-17.78***	-38.89***	-32.35***	-76.20***
	(-0.25)	(-9.75)	(-4.71)	(0.10)	(-8.27)	(-4.22)	(-5.13)	(-9.11)	(-6.22)
Grade 11	5.627	-9.150	15.34***	5.678	21.54*	9.177	25.39***	20.26**	m
	(1.31)	(-0.62)	(4.83)	(1.07)	(2.26)	(1.19)	(3.94)	(3.04)	
Grade 12	m	m	m	m	m	m	m	31.20	m
								(1.54)	
Grade repetition	-13.50	m	-12.60***	-17.29*	-35.43*	-20.05**	-24.43***	-25.00***	-10.72
	(-0.95)		(-3.76)	(-2.39)	(-2.27)	(-2.87)	(-3.95)	(-5.53)	(-1.26)
Vocational track (ref. general)	-56.43***	-9.668	m	m	-78.76***	-55.62***	m	-62.64***	5.045
	(-8.35)	(-1.78)			(-8.99)	(-7.67)		(-7.46)	(0.49)
Attended preschool	2.028	6.402*	7.171*	17.10**	0.621	23.58**	0.627	0.427	7.402
	(0.73)	(2.20)	(2.16)	(2.83)	(0.18)	(2.61)	(0.27)	(0.17)	(1.66)
Index of OTL: Formal mathematics	27.60***	42.53***	21.94***	23.16***	33.28***	33.76***	3.486	20.68***	34.31***
	(9.81)	(18.95)	(11.04)	(10.67)	(9.62)	(11.39)	(1.69)	(9.51)	(10.43)
Constant	483.0***	372.2***	431.9***	500.5***	468.0***	280.4***	512.6***	542.8***	496.5***
	(6.56)	(6.00)	(5.51)	(7.25)	(6.28)	(3.61)	(7.76)	(8.25)	(7.28)
Between school variance	978.3***	1 071.6***	1 417.1***	1 774.9***	1 594.1***	1 686.6***	999.8***	2 983.6***	1 593.0***
	(24.19)	(45.54)	(41.60)	(60.45)	(57.85)	(48.39)	(24.02)	(60.98)	(47.73)
Within school variance	3 672.4***	3 595.5***	2 958.6***	3 264.6***	3 913.4***	3 566.2***	2 546.2***	2 574.3***	3 099.2***
	(218.58)	(242.73)	(252.30)	(236.25)	(216.48)	(248.62)	(189.23)	(215.98)	(218.45)
% of the between school variance explained	60.63	49.66	56.31	40.57	57.78	41.13	66.86	41.59	58.34
% of the within school variance explained	15.06	19.19	23.44	9.03	11.67	9.51	17.98	18.86	11.64
Number of observations	3 073	3 415	3 777	3 337	3 062	4 374	2 804	3 198	3 285

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295733

[Part 1/2]

Table B.13 **Model 12: Individual and school level features**

Individual features	ARG	BGR	BRA	COL	CRI	IDN	JOR	KAZ	MEX
ESCS*	7.906***	9.880***	5.484***	8.352***	9.126***	7.441**	8.629***	13.18***	5.373***
	(6.29)	(7.35)	(3.78)	(3.43)	(6.00)	(2.87)	(5.60)	(6.41)	(6.30)
ESCS squared	1.128	0.686	0.321	1.031	1.169*	0.725	0.00748	-0.0139	1.017**
	(1.67)	(0.96)	(0.67)	(1.48)	(2.05)	(1.07)	(0.01)	(-0.01)	(3.09)
Gender (ref. male)	-21.00***	-15.95***	-26.50***	-30.93***	-28.98***	-7.189***	-29.82**	-3.727*	-19.72***
	(-9.17)	(-7.40)	(-17.66)	(-15.94)	(-15.97)	(-3.62)	(-3.05)	(-2.09)	(-20.24)
Native speaker	3.215	15.46***	7.929	17.78*	5.222	-12.16***	16.14***	2.691	15.21***
	(0.69)	(3.59)	(1.90)	(2.30)	(0.79)	(-5.65)	(3.96)	(0.58)	(5.23)
Grade 7	-77.25***	-62.10**	-71.60***	-68.60***	m	-40.48***	-83.80*	-21.48	-47.92***
	(-9.93)	(-2.85)	(-18.94)	(-13.54)		(-3.68)	(-2.38)	(-0.95)	(-7.14)
Grade 8	-54.62***	-34.41***	-55.42***	-46.85***	-45.05***	-35.22***	-28.06	-18.51***	-30.21***
	(-12.69)	(-4.03)	(-14.79)	(-12.68)	(-12.95)	(-4.09)	(-1.76)	(-3.93)	(-7.29)
Grade 9	-27.34***	ref.	-30.14***	-31.01***	-21.46***	-13.58	-27.38***	ref.	-7.996*
	(-7.34)		(-15.53)	(-9.16)	(-6.92)	(-1.92)	(-4.19)		(-2.38)
Grade 10	ref.	19.32***	ref.	ref.	ref.	ref.	ref.	18.37***	ref.
		(3.35)						(6.90)	
Grade 11	7.328	-2.819	29.14***	31.02***	21.36***	13.17**	m	37.85	14.64**
	(1.01)	(-0.07)	(8.36)	(11.53)	(6.76)	(2.69)		(1.86)	(2.79)
Grade 12	-5.889	m	m	m	63.42***	59.56***	m	m	7.058
	(-0.47)				(3.80)	(3.51)			(0.27)
Repeated a grade	-13.35***	-30.71***	-13.25***	-1.402	-10.71***	-18.90***	-35.46***	-14.14	-29.95***
	(-3.59)	(-3.94)	(-6.93)	(-0.62)	(-3.70)	(-7.54)	(-6.42)	(-1.81)	(-13.28)
Vocational track (ref. general)	22.51**	16.79*	-28.17*	3.747	2.472	1.349	m	-15.70	2.835
	(2.85)	(2.45)	(-2.14)	(0.88)	(0.59)	(0.10)		(-1.74)	(1.12)
Attended preschool	27.10***	10.89**	9.462***	2.673	6.931*	2.788	16.84***	0.399	15.02***
	(7.51)	(2.68)	(6.06)	(0.84)	(2.56)	(0.97)	(5.93)	(0.16)	(8.01)
School features	**ARG**	**BGR**	**BRA**	**COL**	**CRI**	**IDN**	**JOR**	**KAZ**	**MEX**
School's mean ESCS	27.82***	46.74***	39.72***	34.58***	24.58***	20.63**	35.59***	28.67**	23.08***
	(6.66)	(6.20)	(12.70)	(8.09)	(5.15)	(2.82)	(4.65)	(2.90)	(10.96)
Average absentism teachers	-2.675	11.52	-5.261	6.210	-3.268	-0.771	-3.423	-3.648	-5.123
	(-0.72)	(1.82)	(-1.81)	(1.17)	(-0.70)	(-0.05)	(-0.75)	(-0.60)	(-1.71)
Teacher shortage	-0.509	4.640	-0.505	-0.461	-0.179	-5.387	-0.606	3.018	0.773
	(-0.26)	(0.85)	(-0.33)	(-0.22)	(-0.08)	(-1.24)	(-0.29)	(1.04)	(0.65)
% of teachers with professional development	-0.0157	0.203*	0.0286	-0.00231	0.0635	0.0468	0.0147	-0.106	0.00250
	(-0.29)	(2.49)	(0.84)	(-0.03)	(1.24)	(0.50)	(0.27)	(-1.57)	(0.09)
Learning time in minutes	0.117***	0.187*	-0.0 119	0.0 387	-0.106	0.196***	-0.0 228	0.163***	0.0 503*
	(4.07)	(2.47)	(-0.34)	(1.07)	(-1.24)	(3.37)	(-0.28)	(3.54)	(2.52)
Infrastructure	0.195	-0.639	-1.639	2.216	0.704	-0.764	-2.139	0.203	1.268
	(0.10)	(-0.22)	(-1.03)	(0.75)	(0.27)	(-0.17)	(-0.90)	(0.06)	(0.93)
Educational resources	-1.861	2.563	4.684**	0.906	4.450*	4.078	-0.136	2.341	1.233
	(-0.99)	(0.68)	(2.91)	(0.32)	(2.07)	(1.10)	(-0.04)	(0.65)	(0.88)
Disciplinary climate	12.15*	21.72**	16.83***	15.79*	20.79**	17.05	33.47***	17.68*	13.93***
	(2.22)	(2.71)	(4.45)	(2.08)	(3.07)	(1.64)	(4.73)	(2.12)	(5.13)
Absentism pupils	-45.49**	-77.35***	-15.25	-16.49	-4.145	-43.03	-52.80**	-4.415	-65.42***
	(-3.20)	(-4.16)	(-1.50)	(-0.47)	(-0.20)	(-1.39)	(-2.95)	(-0.20)	(-7.55)
% of girls at school	50.94*	12.55	-12.45	17.15	59.57	-42.86*	44.03***	10.50	9.009
	(2.29)	(0.80)	(-0.58)	(1.15)	(1.71)	(-2.04)	(3.93)	(0.46)	(0.82)
Private school (ref. public)	9.165	2.480	-2.440	-2.556	3.568	-0.181	21.85***	-9.658	-8.044*
	(1.74)	(0.11)	(-0.38)	(-0.39)	(0.29)	(-0.02)	(3.34)	(-0.72)	(-2.02)
School size	-0.00605	0.0287**	0.00575*	-0.000564	0.00459	0.0181	-0.00594	-0.0196***	0.00410**
	(-0.94)	(3.01)	(2.18)	(-0.34)	(1.02)	(1.63)	(-0.62)	(-4.40)	(3.03)
Location of school: village (ref. town)	-3.591	31.13	21.05***	14.10	3.642	5.400	15.32*	-6.695	1.447
	(-0.37)	(1.66)	(3.44)	(1.56)	(0.62)	(0.69)	(2.14)	(-0.68)	(0.37)
Location of school: city	4.964	2.037	-10.42**	-2.389	-3.826	-0.293	10.62*	5.785	0.697
	(1.26)	(0.31)	(-3.06)	(-0.43)	(-0.56)	(-0.03)	(2.30)	(0.75)	(0.25)
Constant	379.3***	294.7***	595.4***	568.1***	552.8***	541.5***	206.8***	397.8***	531.9***
	(6.25)	(5.27)	(15.35)	(11.33)	(7.53)	(8.46)	(3.48)	(6.82)	(14.29)
Between school variance	443.3***	788.3***	597.3***	602.9***	496.0***	1 142.7***	582.9***	1 036.9***	573.0***
	(41.57)	(38.55)	(66.87)	(43.31)	(43.72)	(53.36)	(36.01)	(35.32)	(87.43)
Within school variance	2 651.2***	3 865.9***	2 713.8***	2 667.7***	2 082.0***	2 313.6***	3 415.8***	3 043.7***	3 298.6***
	(251.02)	(325.96)	(354.33)	(245.68)	(247.29)	(259.04)	(317.72)	(303.00)	(563.88)
% of the between school variance explained	82.93	83.04	78.01	69.11	75.00	57.13	73.08	44.26	70.44
% of the within school variance explained	18.50	7.08	20.98	26.26	22.87	5.81	11.33	5.90	7.81
Number of observations	5 240	4 916	17 622	7 717	4 192	4 990	5 858	5 751	31 630

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
c: There are too few observations or no observation to provide reliable estimates (i.e. there are fewer than 30 students or fewer than 5 schools with valid data).
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
ref.: reference category.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295743

[Part 2/2]

Table B.13 **Model 12: Individual and school level features**

Individual features	MNE	MYS	PER	ROU	SRB	THA	TUN	TUR	VNM
ESCS*	8.936***	17.76***	9.030***	17.86***	6.310***	15.63***	10.16***	1.211	10.27***
	(5.10)	(8.35)	(5.23)	(9.64)	(4.22)	(7.95)	(5.48)	(0.66)	(3.97)
ESCS squared	-0.649	2.684**	1.341*	2.060*	-1.586	3.005***	2.291***	-1.438*	1.404
	(-0.55)	(3.24)	(2.44)	(2.42)	(-1.13)	(4.12)	(3.70)	(-2.19)	(1.87)
Gender (ref. male)	-13.55***	1.229	-30.65***	-13.06***	-25.95***	-2.220	-26.44***	-27.01***	-24.81***
	(-4.09)	(0.58)	(-15.54)	(-5.38)	(-9.91)	(-1.08)	(-13.17)	(-14.89)	(-12.22)
Native speaker	11.65	-9.802**	19.33***	6.113	3.248	-11.36***	10.11	-0.217	9.576
	(0.74)	(-2.92)	(5.45)	(0.72)	(0.54)	(-3.53)	(1.24)	(-0.05)	(1.30)
Grade 7	m	m	-68.27***	30.84	m	4.117	-90.21***	-14.44	-132.1***
			(-8.69)	(1.29)		(0.15)	(-12.37)	(-0.62)	(-8.59)
Grade 8	28.15	-157.9***	-46.41***	11.60	-60.46*	-18.13	-51.13***	-15.34	-79.17***
	(0.89)	(-3.50)	(-12.01)	(0.99)	(-2.10)	(-1.12)	(-7.57)	(-0.88)	(-6.84)
Grade 9	ref.	-87.71***	-23.97***	ref.	ref.	-17.26***	-18.67**	-34.62***	-55.04***
		(-12.93)	(-8.35)			(-4.55)	(-2.62)	(-11.41)	(-5.16)
Grade 10	12.94***	ref.	ref.	10.78*	33.05***	ref.	ref.	ref.	ref.
	(3.39)			(2.38)	(3.32)				
Grade 11	m	11.80	23.29***	m	m	17.75*	26.81***	20.99**	0
		(0.74)	(9.12)			(2.38)	(5.13)	(3.18)	c
Grade 12	m	m	m	m	m	m	m	29.13	0
								(1.95)	c
Repeated a grade	-32.69***	m	-11.03***	-15.72**	-39.28**	-22.00***	-21.32***	-23.20***	-9.499
	(-3.82)		(-4.03)	(-2.62)	(-3.23)	(-3.62)	(-4.31)	(-6.61)	(-1.24)
Vocational track (ref. general)	-58.63***	-13.63**	m	m	-44.43***	-29.73	m	4.378	9.521
	(-6.99)	(-2.65)			(-5.90)	(-1.60)		(0.57)	(0.45)
Attended preschool	-0.173	10.23***	7.205*	12.87**	-1.604	31.77***	0.508	-0.455	9.188**
	(-0.08)	(4.31)	(2.55)	(2.73)	(-0.53)	(3.87)	(0.21)	(-0.21)	(2.59)
School features	**MNE**	**MYS**	**PER**	**ROU**	**SRB**	**THA**	**TUN**	**TUR**	**VNM**
School's mean ESCS	66.54***	29.90***	33.84***	38.34***	46.55***	19.16**	30.33***	47.44***	34.06***
	(5.62)	(5.58)	(8.16)	(6.06)	(4.15)	(3.02)	(5.43)	(5.50)	(5.03)
Average absentism teachers	12.38	-1.761	-2.056	-5.132	19.05*	-8.347	1.128	5.213	14.36
	(1.51)	(-0.30)	(-0.44)	(-0.59)	(2.03)	(-1.27)	(0.25)	(0.59)	(1.01)
Teacher shortage	1.551	3.260	0.870	-5.823	-3.147	6.259	-5.783*	0.373	-0.0551
	(0.24)	(1.00)	(0.35)	(-1.76)	(-0.77)	(1.90)	(-2.12)	(0.12)	(-0.02)
% of teachers with professional development	0.127	0.112	0.0608	-0.0522	-0.0283	0.0339	-0.00720	0.0477	0.0296
	(1.61)	(1.86)	(1.05)	(-0.94)	(-0.45)	(0.55)	(-0.11)	(0.51)	(0.53)
Learning time in minutes	-0.126	0.116*	-0.00468	0.255**	0.0930	0.130	-0.0510	0.531***	-0.0929
	(-0.70)	(2.35)	(-0.15)	(2.93)	(0.39)	(1.51)	(-0.95)	(4.97)	(-1.67)
Infrastructure	9.543*	0.854	-1.207	-8.035*	-1.144	5.649*	-4.478	-4.310	-0.968
	(2.38)	(0.32)	(-0.52)	(-2.07)	(-0.35)	(1.97)	(-1.55)	(-1.05)	(-0.27)
Educational resources	-7.286	3.183	2.472	11.03**	4.121	2.358	-3.073	5.113	-2.532
	(-1.39)	(1.09)	(1.15)	(3.03)	(1.09)	(0.67)	(-1.05)	(1.23)	(-0.71)
Disciplinary climate	-10.70	35.59***	13.12	22.18***	15.80	17.56	37.09***	58.47***	41.71***
	(-0.87)	(3.91)	(1.72)	(3.37)	(1.72)	(1.21)	(3.31)	(6.24)	(3.73)
Absentism pupils	9.218	-28.59	-79.22***	-45.05**	-178.0***	-41.94	-42.46*	17.18	-49.41
	(0.23)	(-1.57)	(-3.43)	(-3.01)	(-5.13)	(-1.39)	(-2.28)	(0.72)	(-1.42)
% of girls at school	5.642	8.264	10.81	61.29**	20.34	6.366	71.29	9.007	183.0***
	(0.25)	(0.35)	(1.12)	(3.04)	(1.15)	(0.36)	(1.47)	(0.83)	(4.17)
Private school (ref. public)	-93.51***	21.37	3.150	5.716	-92.06***	-24.29***	13.84	0	-30.82*
	(-4.76)	(1.88)	(0.42)	(0.40)	(-5.93)	(-3.70)	(0.62)	c	(-2.37)
School size	-0.00243	0.00558	-0.00343	0.0192***	0.0247**	0.00797*	0.00347	-0.00518	0.00454
	(-0.17)	(1.19)	(-0.74)	(3.79)	(3.03)	(2.55)	(0.51)	(-0.90)	(0.87)
Location of school: village (ref. town)	0	-5.273	-10.97	8.819	0	7.469	-27.24	11.22	-2.475
	c	(-0.91)	(-1.72)	(0.77)	c	(0.73)	(-1.72)	(0.80)	(-0.39)
Location of school: city	15.05	0.411	-0.527	-0.363	5.082	3.190	3.177	-7.806	10.17
	(1.36)	(0.07)	(-0.09)	(-0.07)	(0.77)	(0.53)	(0.48)	(-1.08)	(1.01)
Constant	511.6***	431.5***	548.7***	419.4***	552.8***	267.8***	516.4***	571.6***	477.6***
	(8.22)	(8.01)	(8.53)	(7.12)	(7.13)	(4.00)	(7.71)	(8.55)	(7.13)
Between school variance	296.5***	499.9***	577.9***	698.1***	594.5***	1 062.0***	559.1***	1 051.3***	693.8***
	(17.95)	(41.01)	(45.96)	(45.83)	(31.61)	(28.54)	(32.89)	(44.32)	(38.98)
Within school variance	4 009.9***	3 972.5***	3 089.3***	3 368.3***	4 174.9***	3 751.4***	2 574.8***	2 609.1***	3 264.9***
	(242.89)	(298.01)	(288.47)	(295.20)	(257.82)	(313.41)	(226.77)	(269.62)	(255.28)
% of the between school variance explained	88.07	76.52	82.18	76.63	84.25	62.93	81.47	79.42	81.85
% of the within school variance explained	7.26	10.72	20.06	6.14	5.77	4.81	17.05	17.76	6.92
Number of observations	4 688	5 119	5 603	5 055	3 832	6 351	3 750	4 577	4 811

Notes: Multilevel regression model consists of student and school levels.
Multilevel models are able to analyse data in nested structure (student within school). For this estimation we use random slope model at the school level.
t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
c: There are too few observations or no observation to provide reliable estimates (i.e. there are fewer than 30 students or fewer than 5 schools with valid data).
m: Data are not available. These data were not submitted by the country or were collected but subsequently removed from the publication for technical reasons.
ref.: reference category.
* ESCS refers to the *PISA index of economic, social and cultural status*.
Source: OECD, PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933295743

PISA FOR DEVELOPMENT

Annex C

PISA 2012 mathematics performance and books in the home

Table C.1 PISA 2012 mathematics performance by students' reports of number of books at home

	Number of books				Average score difference (Group 4 – Group 1)
	0-25 (Group 1)	26-100 (Group 2)	101-200 (Group 3)	>201 (Group 4)	
Argentina	372	411	428	436	64
Bulgaria	398	461	483	498	101
Brazil	380	412	426	431	51
Colombia	365	402	415	440	75
Costa Rica	396	431	451	474	78
Indonesia	366	391	393	395	29
Jordan	378	404	419	408	30
Kazakhstan	415	443	455	456	40
Mexico	406	434	441	452	46
Montenegro	374	420	441	455	81
Malaysia	395	432	440	454	59
Peru	351	399	429	417	66
Romania	414	457	481	500	86
Serbia	417	469	486	506	89
Thailand	413	438	452	488	75
Tunisia	380	414	426	428	48
Turkey	421	473	487	513	92
Viet Nam	503	519	534	541	38

Source: Authors' calculations based on PISA 2012 Database.
StatLink http://dx.doi.org/10.1787/888933294103

One indicator of student disadvantage is the number of books in the home, one of the OECD measures of student socio-economic background, as it indicates home support for learning (Carnoy and Rothstein, 2013). Students that come from homes with fewer books have lower levels of performance than that of students who come from homes with more books. The differences in performance between students from homes with fewer than 25 books compared with those from homes with more than 200 books is striking, in some cases close to a full standard deviation on the mathematics scale (for Bulgaria, Montenegro, Romania, Serbia and Turkey) and equivalent to a proficiency level or more. The size of these differences is comparable for the OECD countries examined by Carnoy and Rothstein (2013). Other middle-income countries also show substantial differences (Colombia, Costa Rica and Thailand stand out in this regard). For many other middle-income countries (Indonesia, Jordan, Kazakhstan, and Viet Nam), the difference is lower.

References

Carnoy, M. and **R. Rothstein** (2013), "What do international tests really show about US student performance", Economic Policy Institute, www.epi.org/publication/us-student-performance-testing/.

PISA FOR DEVELOPMENT

Annex D

Summary descriptions for the six levels of proficiency in mathematics

ANNEX D: SUMMARY DESCRIPTIONS FOR THE SIX LEVELS OF PROFICIENCY IN MATHEMATICS

Table D.1 **Summary descriptions for the six levels of proficiency in mathematics**

Level	Lower score limit	Percentage of students able to perform tasks at each level or above (OECD average)	What students can typically do
6	669	3.3%	At Level 6, students can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop new approaches and strategies for attacking novel situations. Students at this level can reflect on their actions, and can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situation.
5	607	12.6%	At Level 5, students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They begin to reflect on their work and can formulate and communicate their interpretations and reasoning.
4	545	30.8%	At Level 4, students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Students at this level can utilise their limited range of skills and can reason with some insight, in straightforward contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments, and actions.
3	482	54.5%	At Level 3, students can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning.
2	420	77.0%	At Level 2, students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions to solve problems involving whole numbers. They are capable of making literal interpretations of the results.
1	358	92.0%	At Level 1, students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are almost always obvious and follow immediately from the given stimuli.

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

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(98 2015 03 1 P1) ISBN 978-92-64-24618-8 – 2015

PISA

The Experience of Middle-Income Countries Participating in PISA 2000-2015

This report provides a systematic review and empirical evidence related to the experiences of middle-income countries and economies participating in the Programme for International Student Assessment (PISA), 2000 to 2015. PISA is a triennial survey that aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. To date, students representing more than 70 countries and economies have participated in the assessment, including 44 middle-income countries, many of which are developing countries receiving foreign aid. This report provides answers to six important questions about these middle-income countries and their experiences of participating in PISA: What is the extent of developing country participation in PISA and other international learning assessments? Why do these countries join PISA? What are the financial, technical, and cultural challenges for their participation in PISA? What impact has participation had on their national assessment capacity? How have PISA results influenced their national policy discussions? And what does PISA data tell us about education in these countries and the policies and practices that influence student performance?

The findings of this report are being used by the OECD to support its efforts to make PISA more relevant to a wider range of countries, and by the World Bank as part of its on-going dialogue with its client countries regarding participation in international large-scale assessments.

Contents

Consult this publication on line at: ***http://dx.doi.org/10.1787/9789264246195-en***

This work is published on the *OECD iLibrary,* which gathers all OECD books, periodicals and statistical databases.
Visit ***www.oecd-ilibrary.org*** and do not hesitate to contact us for more information.

2015

OECD*publishing*
www.oecd.org/publishing

READ
Russia Education Aid for Development
When children learn, nations prosper

ISBN 978-92-64-24618-8
98 2015 03 1P

9 789264 246188