PISA for Development
Project Completion Report

PISA for Development International Technical Workshop
3 December 2020
Paris, France (Virtual event)

OECD

Table of contents

Tables

Figures

Boxes

Abbreviations and acronyms

ASER	Annual Status of Education Report
CONFEMEN	Conference of the Ministers of Education of French-speaking countries
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
ESCS	Economic, social and cultural status
ETS	Educational Testing Service
IAG	International Advisory Group
LAMP	Literacy Assessment and Monitoring Programme
LLECE	Latin American Laboratory for Assessment of the Quality of Education
LMIC	Low and Middle Income Countries
NPM	National Project Manager
OECD	Organisation for Economic Co-operation and Development
PASEC	Programme for the Analysis of Education Systems
PBG	PISA Governing Board
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Students Assessment
PISA-D	PISA for Development
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SDG	Sustainable Development Goal
STEP	Skills Toward Employment and Productivity
TAG	Technical Advisory Group
TIMSS	Trends in International Mathematics and Science Study
UIS	UNESCO Institute of Statistics
UNESCO	United Nation Educational, Scientific and Cultural Organisation
UNICEF	United Nations International Children's Emergency Fund
WEI-SPS	World Education Indicators-Survey of Primary Schools

1. Introduction

As per the PISA for Development (PISA-D) project design, this Project Completion Report (PCR) has been prepared for a wider audience that consolidates the evidence from the three technical strands of work (A, B and C), the policy relevance for participating countries, lessons learnt, best practices and main challenges related to the implementation of PISA in low-and-middle-income-countries (LMIC). A preliminary version of the report was published in 2019 as planned and was presented at an international seminar with participating countries, other PISA participants, development partners and technical partners held in London on 25 September 2019. This final version of the report has been prepared following the completion of Strand C of the project and the release of the final outputs of the initiative in December 2020. This final version of the report is being presented at a virtual *International Technical Workshop* convened by the OECD from its headquarters in Paris on 3 December 2020.

This report should be read in conjunction with the following key documents and reports that have been produced as part of the PISA-D initiative:

> PISA for Development Technical Report - Strand A and B
>
> PISA for Development Technical Report – Strand C
>
> Technical Papers for Strands A, B and C
>
> PISA for Development Assessment and Analytical Framework
>
> The Experience of Middle-Income-Countries Participating in PISA: 2000-2015
>
> A Review of International Large-scale Assessments
>
> Making Education Count for Development
>
> PISA for Development Capacity Needs and Capacity Building Reports
>
> PISA for Development National Reports
>
> PISA for Development Results in Focus Nos. 91 and 106

In addition, the PISA for Development database (Strands A and B) is available at this link and the database for Strand C is available at this xlink.

This report is structured in accordance with the agreed PISA for Development logical framework (Annex A) and focuses primarily on the project's five main outputs, which are:

- Descriptive power of cognitive assessments enhanced (reading, mathematics and science).

- Contextual questionnaires and data-collection instruments enhanced (system-level, for students, parents and schools).

- An analytical framework and methodological approach for including out-of-school 15-year-olds in assessments developed.

- Country capacity in assessment and analysis strengthened among participants.

- Engagement established with LMIC and partners for peer-to-peer analysis and learning opportunities to support the United Nations (UN)-led post-2015 process,

later known as the Sustainable Development Goals (SDGs) and also known as the Global Goals and the 2030 Agenda for Sustainable Development.

1.1. PISA and PISA for Development

Over the past two decades, the OECD Programme for International Student Assessment (PISA) has become the world's premier reference for evaluating the quality, equity and efficiency of school systems. By identifying the characteristics of high-performing and improving education systems, PISA allows governments and educators to identify effective policies that they can adapt to their local contexts. PISA is now being used by the UN system as a major source of data for monitoring progress towards the Sustainable Development Goals (SDGs) established by the international community as the blueprint to achieve a better and more sustainable future for all.

PISA for Development (PISA-D), the latest of the OECD's PISA assessments, focused on making PISA more accessible and relevant to LMIC. Through this initiative the OECD has enhanced its PISA instruments so that they target the range of student performance in these countries more effectively. As part of this initiative, the OECD has also collected background information to capture how students learn, teachers teach and schools operate in these contexts. In addition, PISA-D has, for the first time in the history of international large-scale assessments, included out-of-school youth of school-age in its coverage. PISA-D has also helped the nine participating countries to build their capacity to manage large-scale assessments and to make use of the results in support of national policy dialogue and education policy-making.

1.1.1. The experience of middle-income countries in PISA

Students representing more than 80 countries and economies that together make up over 80% of the world economy have participated in PISA since its launch, including 44 middle-income countries, 27 of which have been recipients of foreign aid. As more and more participants join it has become apparent that the design and implementation models for PISA need to evolve to successfully cater to a larger and more diverse set of countries, including the growing number of LMIC who want to participate in the assessment (Lockheed, Prokic-Bruer and Shadrova, 2015[1]). In particular, PISA needs to take more account of the marked differences between high- and middle-income countries in education quality and equity and their correlates.

The OECD's analysis of the experience of middle-income countries in PISA revealed three key results that have implications for the further development of the assessment:

- First, the overall performance of 15-year-old students in all the middle-income countries participating in PISA, except Viet Nam, is lower than that of students in OECD countries, and varies widely. Performance is also concentrated at the lower levels of the PISA proficiency scales.

- Second, some of the educational inputs as currently measured by PISA are unrelated to differences in performance across schools in the majority of the middle-income countries that participate in PISA. Also, the measure of economic, social and cultural status (ESCS) used by PISA does not adequately capture lower levels of parental education, income and risk factors of poverty that are more frequent in LMIC. The data captured on the context surrounding students could be made more relevant, particularly in respect of policies.

- Third, out-of-school rates for lower secondary school children are high in many LMIC and, in addition, many 15-year-olds in these contexts are also enrolled in grades below those that are eligible for PISA (i.e. grade 6 and below). The combination of these two exclusion mechanisms result in indices as low as 50% coverage of the 15-year-old population in some PISA-participating countries, and limit the comparability of middle-income countries' results with other countries. Unless the assessment takes concrete steps to incorporate all the 15-year-olds in a country's population in the survey, PISA runs the risk of reinforcing policies of exclusion in middle-income countries.

1.1.2. The PISA for Development initiative

Building on the experience of middle-income countries in PISA since 2000, and in an effort to respond to the emerging demand for PISA to cater for a wider range of countries, the OECD launched the PISA for Development (PISA-D) initiative in 2014. This unique pilot project spanning six years aimed to make the assessment more accessible and relevant to a wider range of countries by:

- increasing the resolution of the PISA tests at the lower end of the student performance distribution (technical strand A)

- capturing a wider range of social and economic contexts (technical strand B)

- incorporating an assessment of out-of-school 14-16 year-olds (technical strand C).

PISA-D has also been informed by analysis of the lessons and experiences from other regional and international large-scale assessments in education in LMIC (Cresswell, Schwantner and Waters, 2015[2]). The project contributes to the monitoring of international educational targets related to the Education Sustainable Development Goal (SDG), adopted by the United Nations General Assembly in 2015 as part of the Agenda for Sustainable Development.

Eight countries participated in the PISA-D school-based assessment: Bhutan[1], Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia. Participation in the project was by demand; the OECD did not reject any countries that expressed an interest in participating. One of the main reasons for the countries' participation is policy makers' wish to understand why students in their countries achieve certain levels of performance. Assessment results provide these policy makers with data and evidence that can be used to determine what they can do to improve their educational systems and, ultimately, ensure that their students obtain the skills needed to succeed in tomorrow's world and as set out in the education SDG framework.

Six countries participated in the PISA-D out-of-school assessment: Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia. However, of these six participants, only Guatemala, Honduras, Panama, Paraguay and Senegal completed the Main Survey. As with participation in Strands A and B of the project, participation was by demand. PISA-D's definition of out-of-school youth incorporates all those 14-16 year-olds that are not reflected in the school-based survey, including those who are out-of-school and those who are in school but enrolled at grade 6 or below. The out-of-school component adopts the same framework used for the school-based component, as the description of competencies,

---

[1] Bhutan did not participate fully (only in a Pilot Field Trial) and the country's data are not included in the international PISA-D database

particularly at lower levels of performance, also apply to the out-of-school population. While the framework is the same, the out-of-school instruments had to be shorter and thus they focus even more on the lower levels of performance and do not include science. Through the out-of-school assessment, PISA-D is able to report on what all 15-year-olds in a population know and can do. The analysis of these data yield valuable insights for governments, in particular about the effectiveness of their education systems, and about the success of policies that aim to ensure inclusive and equitable quality education and learning opportunities for all. It will also serve to reinforce policies of inclusion and contribute to the monitoring and achievement of the education SDG with its emphasis on leaving no one behind.

The PISA-D in-school instruments have been available for use in PISA from the 2022 cycle onwards and allow LMIC to participate in PISA more meaningfully. The PISA-D out-of-school instruments are being made available for use in PISA from the 2025 cycle onwards and are also being developed for use as part of multi-sector household surveys. The enhanced instruments also support global measures of reading and mathematical skills as part of the education SDG agenda, strengthening PISA's potential to provide a global metric for measuring progress towards the education SDG targets and indicators. These enhancements of the instrumentation are not just important for LMIC, they are of benefit to the PISA programme as a whole.

1.1.3. What PISA-D adds to PISA

While PISA-D was implemented within the overall PISA framework and in accordance with PISA's technical standards and usual practices, new features and enhancements made the assessment more accessible and relevant to LMIC:

- An equal treatment of the major domains tested: reading, mathematics and science – unlike PISA, where one of the domains is given a particular focus in each cycle.

- Targeted test instruments that cover a wider range of performance at the lower levels of proficiency, while still providing scores that cover the whole of the PISA framework and are comparable to the main PISA results – unlike PISA where the tests are not targeted on particular levels of performance.

- Modified test instruments and questionnaires with a reduced reading burden, recognising lower reading literacy capacity in LMIC.

- Contextual questionnaires that have at their core items from PISA to facilitate international comparisons, but also include several distinct PISA-D items that are more relevant to LMIC. These new items also respond to the policy priorities of the countries participating in PISA-D.

- An assessment of the out-of-school population: PISA assesses 15-year-olds in grade 7 or above. PISA-D assesses this same population but also assesses 14-16 year-olds who are not in school or are in grade 6 or below.

These enhancements are not just valuable for LMIC, OECD member countries benefit from the greater granularity now available for measuring performance at the lower levels – 20% of students in OECD member countries perform at these levels. In addition, some OECD member countries, such as Mexico and Turkey, have significant proportions of their 15-year-olds out-of-school and so may benefit from the methodologies and approaches for reaching this most disadvantaged population piloted in PISA-D.

Another feature unique to PISA-D is the learning and capacity-building opportunities built into each phase of project implementation. In preparing to implement the assessment, PISA-D countries underwent a capacity needs analysis based on PISA's technical standards and devised a capacity-building plan that is also relevant for strengthening their national assessment systems. Countries also prepared a project implementation plan, guiding their implementation of the survey and ensuring that the necessary human and financial resources are in place. While PISA countries have not benefitted from similar support in the past, the PISA-D project served as the basis for developing a model of optional support within the core PISA survey, which has been offered more widely to all new participating countries from the 2022 cycle onwards – four countries have taken up this option, namely: El Salvador, India, Mongolia and Uzbekistan.

The learning and capacity building opportunities in PISA-D also included support for data analysis and reporting. PISA-D results for Strands A, B and C were published in national reports produced by the countries in collaboration with the OECD. It was one of the significant achievements of PISA-D that participating countries demonstrated active engagement in the analysis, reporting and use of the data.

As part of the national report production process, the OECD and its contractors provided inputs to the countries to strengthen their capacities for data analysis, interpretation of PISA results, report writing and the production of tailored communication products to support the dissemination of PISA results and policy messages. These reports and other communication products present results in the context of the international PISA scales and include relevant analyses and information based on the policy priorities of each country. The reports constitute a summary of key results and analysis designed to stimulate a constructive debate on improvement, building upon and enriching already existing data and evidence from national, regional or international sources. The reports were the culmination of an engagement and communication strategy implemented by each country – another new feature introduced by PISA-D – which involved key stakeholders in the discussion of the results and policy implications. Stakeholders include pupils, parents, teachers, teacher unions, school principals, academia, civil society, media, and central and local government.

The PISA-D modality for supporting data analysis and reporting has also been mainstreamed in main PISA. Seven countries (Albania, Bosnia-Herzegovina, Indonesia, Moldova, Panama, Ukraine and Serbia) participated successfully in OECD's PISA 2018 Lead Analysts Programme. As with the PISA-D programme, the OECD and its contractors provided inputs to the seven countries participating in the PISA 2018 Lead Analysts Programme to strengthen their capacities for data analysis, interpretation of PISA results, report writing and the production of tailored communication products to support the dissemination of PISA results and policy messages. Six countries (El Salvador, India, Jamaica, Mongolia, Romania and Uzbekistan) have already signed up for the OECD's PISA 2022 Lead Analysts Programme and it is expected that at least four more will join the programme before it commences in 2023.

1.2. Key features of PISA-D

1.2.1. Content

- The school-based survey assessed reading, mathematics and science. The PISA-D household-based survey assessed 14-16 year-olds not enrolled in PISA's target grades (grade 7 and above) in the domains of reading and mathematics. In the PISA-D surveys (school-based and household-based), each domain was treated

equally in the assessment. PISA-D assesses not only whether students and out-of-school youth can reproduce knowledge, but also whether they can extrapolate from what they have learnt and apply their knowledge in new situations. It emphasises the mastery of processes, the understanding of concepts, and the ability to function in various types of situations.

1.2.2. Participating students and out-of-school youth

- Around 37 000 students completed the school-based assessment, representing about a million 15-year-old students (in grade 7 or above) in the schools of the participating countries. More than 7 200 respondents completed the household-based assessment, representing 1 431 497 14-16 year-olds who were either not in school or were in school in grade six or below across five countries: Guatemala, Honduras, Panama, Paraguay and Senegal[2].

1.2.3. The assessment

- In the **school-based assessment**, paper-based tests were used and each student received 60 minutes of assessment time for each of two cognitive domains.

- Test items were a mixture of multiple-choice questions and questions requiring students to construct their own responses. The items were organised in groups or "units"; each unit was based on a stimulus (e.g. a graph, short paragraph, or passage) that sets out a real-life situation. The school-based assessment draws on about 195 test items across all domains.

- More than half of the assessment items were identical to the items used PISA 2015 paper-based assessment, enabling the reporting of results on the PISA scale through scale-linking methods. The remaining items comprised adapted PISA items (e.g. with extended scoring rubrics) and items used in other OECD skills assessments, which were evaluated against PISA frameworks.

- The tests were targeted at the lower levels of performance as measured in the PISA scale. In the PISA-D test, more than two-thirds of the items are at Level 2 and below; and less than one-third at Levels 3 through 6. Level 2 marks the level of proficiency at which students begin to demonstrate the competencies that will enable them to participate effectively and productively in life as continuing students, workers and citizens. For the measurement of the education SDG, Level 2 is taken by the UN system as the minimum level of proficiency in reading and mathematics that all children should achieve at the end of lower secondary school.

- Students also answered a background questionnaire, which took 35 minutes to complete. The questionnaire seeks information about the students themselves, their well-being, educational attainment and engagement, their homes, their families, and their school and learning experiences. School principals completed a school questionnaire that describes the school, its students and teachers, and the learning environment. Teachers completed a questionnaire about themselves, the school's resources, their teaching practice and their students.

---

[2] While Zambia participated in the household survey component of PISA-D, it did not complete a Main Survey data collection.

- In the **household-based assessment**, the background interviews and cognitive assessments were tablet-based, took place mainly in households, and lasted a little more than 90 minutes for each respondent.

- An interviewer conducted the household-based assessment beginning with a 30-minute interview during which the respondent answered a series of background questions covering topics such as the youth's school and learning experiences, well-being and home life. The interviewer recorded the responses on the tablet.

- The cognitive assessments began with a 10-minute core module composed of five reading and five mathematics items. Depending upon the responses to these, the youth was guided automatically to either the full 45-minute cognitive test (approximately 32 reading and mathematics literacy items) or a 30-minute assessment of reading components (sentence comprehension) designed to paint a more nuanced picture of low performance.

- The cognitive assessments were in the languages of instruction used in the participating countries' schools and relied on automatically scored items only.

- The tests used mostly a subset of items from the in-school assessment of reading and mathematics, which allowed for reporting results on the PISA scale through scale-linking methods.

- The items were targeted at the lower levels of performance as measured on the PISA scale. In the out-of-school assessment, item selection focused on the scale at or below Level 2 with an emphasis on the lower end of the scale. As in the in-school assessment, coverage of all processes was maintained and contexts of the items were reviewed to ensure appropriateness for what individuals would encounter in an out-of-school context.

- Parents (or the person most knowledgeable about the young person) also answered a paper-based questionnaire about the youth's background and childhood experiences.

- The interviewer completed a short household-observation module on the tablet, which included questions about the location of the household, aspects of the neighbourhood, and some characteristics of the dwelling.

1.3. The contextual framework for PISA-D

The focus of the PISA contextual questionnaires is on understanding how measures of student performance at age 15 are related to various aspects of school and classroom practice as well as other related factors, such as economic, social and cultural context. The PISA-D questionnaires include these aspects and cover a broader set of well-being outcomes and a wider range of risk and protective factors, taking into account differences in life experiences of children in LMIC.

The PISA-D questionnaire framework uses the Education Prosperity model (Willms, 2015[3]) as an overarching framework, while also taking into account the goals of PISA-D, lessons from past PISA cycles and other international studies, recommendations from research literature and the priorities of the participating countries. Education prosperity is a life-course approach that includes a core set of metrics for success at six key stages of development, covering the period from conception to adolescence. It identifies a key set of outcomes called "Prosperity Outcomes" for six stages of development from conception to

age 18, and a set of family, institutional and community factors, called "Foundations for Success", which drive these outcomes. PISA-D focuses on the fifth stage of the Educational Prosperity framework, late primary and lower secondary (ages 10 to 15).

The framework places great emphasis on equality and equity, with equality referring to differences among sub-populations in the distribution of their educational outcomes and equity referring to differences among sub-populations in their access to the resources and schooling processes that affect schooling outcomes. The framework also focuses on the measurement of socio-economic status and poverty, extending the measure of the PISA index of ESCS and exploring an international measure of poverty for youth in LMIC.

The PISA-D questionnaire framework focuses on 15 modules of content, which measure the four Prosperity Outcomes, the five Foundations for Success, and the six demographic factors relevant to assessing equality and equity. The questionnaires also include teacher, school and system-level background measures that provide additional context.

Table 1.1. Modules assessed in the PISA-D questionnaires

1. Prosperity Outcomes	1.1 Academic performance (measured through the PISA-D tests)
	1.2 Educational attainment
	1.3 Health and well-being
	1.4 Attitudes towards school and learning
2. Foundation of Success	2.1 Inclusive environments
	2.2 Quality instruction
	2.3 Learning time
	2.4 Material resources
	2.5 Family and community support
3. Demographic factors for assessing equality and equity	3.1 Gender
	3.2 Socio-economic status and poverty
	3.3 Language spoken at home and language of instruction
	3.4 Urban/rural status
	3.5 Immigrant status
	3.6 Disability

PISA-D enhanced the contextual questionnaires to better measure factors that are more strongly related to student performance in LMIC, while maintaining comparability with PISA on a set of indicators, e.g. the questionnaires collect more detailed data on students' socio-economic status as measured by home possessions and parents' education, literacy skills and participation in the labour force.

System-level data describing the general structure of the education systems is also used in the PISA-D analysis and country reports, including information on the structure of national programmes, national assessments and examinations, instruction time, teacher training and salaries, educational finance (including enrolment), national accounts and population data. Available data on all of these indicators were reviewed for PISA-D countries, identifying the current status of system-level data collection and availability in terms of quality and completeness (UNESCO Institute for Statistics, 2016[4]).

1.4. Achievement of the PISA-D objectives

With the completion of the PISA-D project at the end of 2020, this report looks at the extent to which the project's objectives have been achieved. A summary of the achievements follows.

1.4.1 Enhance PISA to make it more relevant to a wider range of countries and thus enable greater PISA participation by LMIC

The OECD has enhanced its PISA instruments through the PISA-D initiative to provide a finer-grained view of low-performing students and to better measure factors more strongly related to student performance in LMIC as described in the *PISA-D Assessment and Analytical Framework* (OECD, 2018[5]).

The field trials of the PISA-D instruments and, most importantly, the Main Survey confirmed that the cognitive instruments have been successfully targeted on the levels of proficiency likely to be found in LMIC; the contextual questionnaires have been enhanced to reflect low-income contexts; and a methodology and approach for incorporating out-of-school 15-year-olds has been designed and piloted. The school-based PISA-D results can be placed on comparable reporting scales to those of the main PISA test

The field trial of the school-based assessment took place from August to December 2016 in Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia; and the field trial of the out-of-school assessment took place from April to September 2017 in Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia. The results informed the main survey design. Main data collection was conducted from August to December 2017 for the school-based assessment and from August to December 2018 for the out-of-school assessment. The PISA-D instruments were mainstreamed into PISA for the 2021 cycle.

All PISA-D countries have affirmed their intention to participate in future cycles of PISA and half of them are participating in PISA 2022. Moreover, the number of LMIC participating in PISA has almost doubled during the period of project implementation, which serves as an additional indicator that PISA-D enabled greater PISA participation by LMIC. However, it is important to note that only one low income country (LIC), Malawi, has formally expressed interest in participating in PISA – this country is preparing for participation in PISA 2025.

1.4.2 Help to build the capacity of participating countries to conduct large-scale learning assessments, and analyse and use the results to support national policies and evidence-based decision making

PISA-D countries prepared for their participation through a process that starts with an analysis of their capacity to implement PISA and make use of PISA data, and includes planning to strengthen that capacity. Countries were supported by the OECD and its contractors at each stage of implementing the assessment cycle. This process helped countries overcome two potential barriers to participation in PISA: a lack of capacity to implement the assessment and a lack of experience in using PISA data and results. To overcome the latter obstacle, the OECD and its contractors offered training and assistance in data analysis, the interpretation of PISA results, report writing and communication.

The focus of the capacity building in the project was on the National Centres and the National Project Managers (NPM) and their teams. The National Centres were all

government institutions and ranged from national institutes of assessment and evaluation, such as INEVAL in Ecuador and INEADE in Senegal, to national examination centres, such as ECZ in Zambia and BCSEA in Bhutan, and Departments and Units of ministries of education, as in Cambodia, Guatemala, Honduras, Paraguay and Panama. In almost all cases, these centres were responsible for all the learning assessments in their respective countries and so the capacity building gained through PISA-D benefited all of the learning assessments in the countries. The NPMs were all government employees who are playing a key role in learning assessment in the countries as managers of surveys and assessments. The NPM's role was to manage PISA-D in the country and be responsible for all aspects of the participating countries' engagement in the programme. The capacities strengthened through PISA-D are therefore relevant to countries' management of their own national assessments and other large-scale international or regional assessments in which they might participate.

Capacity building for implementation was conducted primarily through the PISA-D National Project Manager meetings and other international training events. Ten of these meetings took place during the 2015-18 period and were attended by all country teams. The training was delivered by OECD contractors. Country representatives who attended the meetings were responsible for disseminating the training to the rest of the staff at their National Centres. The OECD and its contractors complemented the international training events with weekly calls, scheduled teleconferences, webinars and country visits.

In addition, the programme followed by each country's lead analyst through the residential programme at the OECD in Paris was a significant capacity building input. The lead analysts were government employees, usually from the National Centre, that were loaned to the OECD for the purpose of collaborating with the OECD over the analysis of the PISA-D data, interpretation of the results and the preparation of a national report. The residential programme at the OECD for the lead analysts included a mix of training, peer learning, mentoring, learning by doing and actual analysis of data and preparation of a national report following a template provided by the OECD. The participating countries also formed analysis task forces at the National Centres to support their Lead Analysts during the period of residence at the OECD in Paris. Through these analysis task forces the capacity building benefits of the programme were spread more widely and ownership of the national reports and key findings was built in each country.

The PISA-D cycle included a comprehensive set of tasks that countries had to complete throughout the duration of the project in accordance with the PISA technical standards. In completing each task, countries demonstrated the necessary capacity to advance to the next step of the work. It is important to note that the country teams in the National Centres in all the countries were government employees and nationals of the countries. In addition, the work on the national report was scheduled in accordance with the timeline for the preparation and launch of the OECD's international data products associated with the project. Each country completed all tasks as scheduled, which indicates that countries developed the capacity needed to administer the assessment, successfully analyse their data, prepare a national report and make full use of the assessment.

1.4.3 Provide policy makers in the participating countries with insights on how to help students learn better, teachers to teach better and school systems to operate more effectively

The field trials of the PISA-D cognitive tests for the in-school assessment provided information about the data collected and survey operations, assessed the quality of the test

items, and helped determine the reliability and comparability of the PISA-D and PISA scales. Using the background questionnaires during the trials allowed for selecting items for the final instruments based on their psychometric properties. PISA's technical standards were applied at every stage of the project. The main survey data collection was subject to a strict adjudication process, particularly for the sampling and translation/adaptation parts of the implementation. The way the sample is selected and the instruments are localised in each country is an integral part of the evidence supporting the construct validity of the intended uses of the assessment.

During the analysis phase of the project, analysts took an in-depth look at the results, including the functionality of all test items, and content experts confirmed that the instruments measure what they purport to measure against the scale descriptions as presented in the PISA-D framework. The analysis validated the assumptions that shaped the PISA-D framework. As a result, participating countries have assessment results that provide relevant data that can inform decisions concerning national policies. Each country's national report highlights main messages from the results as well as policy options to pursue to improve learning outcomes. The national reports were launched in seven of the countries in December 2018 with the full support of the OECD and several of the PISA-D development partners. The countries were aided in their communication of results and supported by the OECD as they launched national policy dialogues based on the reports. Bhutan launched its national report in March 2019 with similar support from the OECD and its development partners and Panama launched its national report (on PISA 2018 and PISA-D Strand C) in December 2019. The four other Strand C countries that completed the Main Survey launched their reports of the out-of-school assessment in December 2020.

1.4.4 Contribute to the monitoring and achievement of the education SDG, which emphasises quality and equity of learning outcomes for children, young people and adults

The education SDG represents a global commitment to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. The worldwide push requires establishing reliable, universal learning metrics on the quality of learning outcomes. PISA and PISA-D have been selected by the UNESCO Institute of Statistics (UIS) and the UN Statistical Commission (the two bodies responsible for monitoring progress towards the education SDG) and accepted by the Education SDG and SDG monitoring global architecture as an internationally comparable measure of the SDG global indicator 4.1.1.c. This is in part due to PISA-D extending the coverage of the assessment to an increasing number of LMIC and including out-of-school youth in the assessment, thus allowing wider measurement of global progress. SDG global indicator 4.1.1.c is:

- Proportion of children and young people at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.

The minimum proficiency level is the benchmark of basic knowledge in a domain (mathematics or reading) measured through learning assessments. In the case of PISA, the reading and mathematics tests have six proficiency levels, of which Level 2 is described as the minimum proficiency level.

With increased numbers of LMIC participating in future cycles of PISA, and with each of these using the results of the assessment to achieve positive changes in teaching and learning, there is global progress towards improving learning outcomes and thus

achievement of the education SDG's aim: all children and young people achieving at least minimum levels of proficiency in reading and mathematics at the end of lower secondary education.

2. Implementation of the PISA for Development project

PISA-D project implementation consisted of five phases over the course of 2014 to 2020.

1. *Design, planning and co-ordination (2014-15)*: Producing expert papers to inform the work of enhancing the assessment instruments, selecting international contractors to conduct the work, and preparing participating countries, including Capacity Needs Analysis and developing a Capacity Building Plan and a Project Implementation Plan for each country. This phase also included the first and second annual meetings of the PISA-D International Advisory Group (IAG) and the first and second annual meetings of the PISA-D Technical Advisory Group (TAG), which were crucial for reaching agreements on the design of the initiative.

2. *Technical development (2015-16)*: Reviewing assessment frameworks and items, selecting items, designing enhancements, preparing materials, and planning for field trials, as well as development of the PISA-D Analysis and Reporting Plan and capacity building for the participating countries.

3. *Field trials and in-country data collection (2016-19)*: Field trials in each country to test the enhanced instruments, reviewing and analysing the results of the field trial, preparing materials for the Main Survey data collection, and conducting the Main Survey data collection as well as capacity building for the participating countries.

4. *Analysis and report writing (2018-20)*: Data cleaning and analysis, interpreting results, countries engaging in and analysing their own data and writing their own national reports supported by OECD and its contractors, including capacity building and training focused on the lead analysts nominated by the participating countries.

5. *Report production, dissemination and post-pilot governance (2018-20)*: Instruments finalised, national reports and international datasets published, a project completion report and technical reports (Strands A&B and Strand C) published, a PISA-D international seminar, and PISA-D instruments incorporated in PISA from the 2021 cycle onwards.

Figure 2.1 shows a detailed breakdown of the activities spanning the six years from 2014 to 2020.

Figure 2.1. PISA for Development implementation schedule
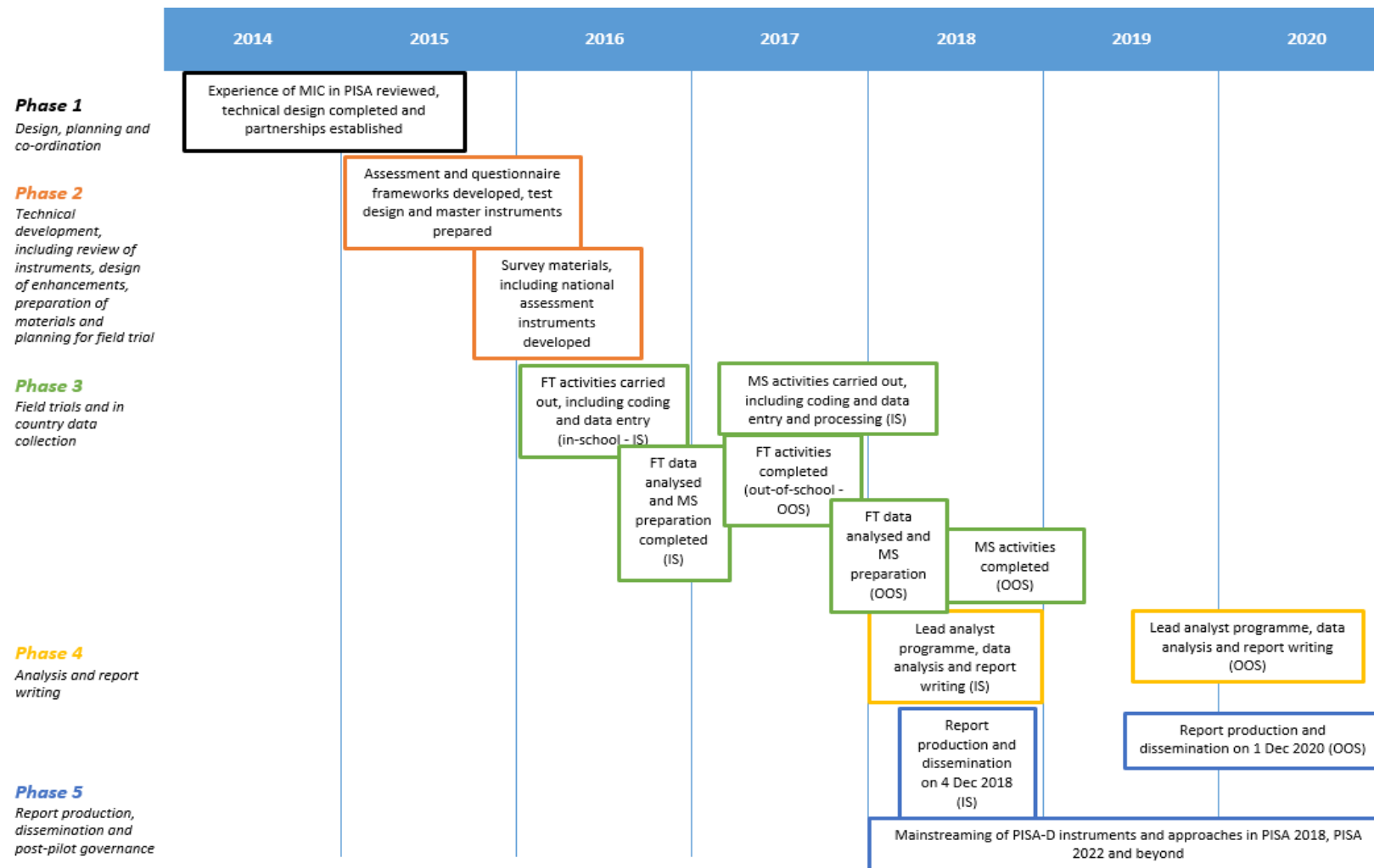
2.1. A collaborative project

PISA-D has been a highly collaborative effort carried out by the OECD, contractors and nine participating countries with support from development partners and technical partners.

The OECD's Directorate for Education and Skills and the Development Co-operation Directorate shared responsibility for the overall management of PISA-D, monitoring its implementation on a day-to-day basis and building consensus among countries. The OECD served as the Secretariat and interlocutor between the PISA-D International Advisory Group (IAG), the PISA Governing Board (PGB), the Technical Advisory Group (TAG) and the PISA-D contractors. The OECD was also responsible for the capacity building of the participating countries, the production of the indicators, the analysis of results, and the preparation of the national reports and project publications in co-operation with the contractors and in close collaboration with the participating countries both at the policy level with the PGB and IAG, at the technical level with the TAG and at the implementation level with the National Project Managers (NPMs).

The PISA-D IAG met annually from 2014 to 2018, comprised of government officials from participating countries, representatives of development partners supporting the initiative, representatives of technical partners, such as UNESCO and UNICEF, invited experts and representatives of the OECD.

The PGB, representing all countries/economies with full PISA membership at senior policy levels, determines the policy priorities for PISA in the context of OECD objectives and oversees adherence to these priorities during the implementation of the programme. The PGB sets priorities for developing indicators, for establishing assessment instruments and for reporting results. Experts from participating countries/economies also serve on working groups to ensure that the instruments are internationally valid and take into account differences in the cultures and education systems.

The PISA-D TAG, managed by the OECD, explored technical issues that have policy or project implications and advised the OECD and its contractors on these issues.

The PISA-D contractors were responsible for survey operations and management and led supporting the countries to implement the programme. The contractors also led development of the enhanced assessment instruments, drawing on the technical expertise of the Subject Matter Expert Groups and Questionnaire Expert Groups that support PISA. The development of the PISA-D frameworks for reading, mathematics and science and the development of the PISA-D cognitive instruments were the responsibility of the contractor Educational Testing Service (ETS), while the design and development of the PISA-D questionnaires were the responsibility of the contractor The Learning Bar. Management and oversight of this survey, the development of the instruments, scaling and analysis were the responsibility of ETS, along with the development of the electronic platform. Other partners or subcontractors involved with ETS include Pearson for the development of the cognitive frameworks, cApStAn for linguistic quality assurance and control and Westat for survey operations and sampling.

As noted above, participating countries implemented the survey at the national level through National Centres (all of which were government agencies or units of the Ministry of Education), managed by NPMs, subject to the agreed administration procedures and in accordance with the PISA-D technical standards put in place by the OECD and its contractors. Also as noted above, the National Centres are also responsible for other learning assessments in the countries. The NPMs play a vital role in ensuring that

implementation is of high quality and help to shape and guide the project in accordance with the PISA-D technical standards. They also verify and evaluate the survey results, analyses, reports and publications. The co-operation of students, teachers and principals in participating schools is crucial to the success of PISA-D during all stages of development and implementation. National experts from the participating countries contributed to the preparation of the frameworks and instruments, and they also provided input for the design of analytical outputs. These national experts were almost all government employees and were mainly drawn from national universities. National Centres collaborated with OECD on the analysis of national data and the production of national reports and other communication products.

From the outset of the project, OECD engaged the participation of the key international agencies and programmes concerned with student assessment and improving the quality of education in LMIC. These technical partners include UNESCO, UIS, the Global Education Monitoring Report team, UNICEF, the Global Partnership for Education and the following assessment programmes: ASER, EGRA, EGMA, SACMEQ, PASEC, Pre-PIRLS and PIRLS, TIMSS, LLECE, STEP, LAMP, Uwezo, and WEI-SPS.[3] Representatives of these agencies and programmes took part in many of the technical workshops and meetings and were invited to comment on all aspects of project design and development.

The international and national costs of the project were funded through a combination of development partner support and financing from the PISA-D countries. The development partners that provided financing or aid-in-kind are France (*Agence française de développement* / French Development Agency); Germany (*Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung* / Federal Ministry for Cooperation and Development and *Deutsche Gesellschaft für Internationale Zusammenarbeit* / German Corporation for International Cooperation); Global Partnership for Education; Inter-American Development Bank; Ireland (Irish Aid); Japan (*独立行政法人国際協力機構* / Japan International Cooperation Agency); Korea; Microsoft Corporation; Porticus Foundation; Positivo; Norway (*Norad er direktoratet for utviklingssamarbeid* / Norwegian Agency for Development Cooperation); Sunny Varkey Foundation; United Kingdom (Department for International Development); and the World Bank.

Most of the development partners and participating countries prepared their support for PISA-D alongside each other and all development partners and participating countries participated together in the international workshops and meetings that were convened by the OECD to prepare the project and to oversee implementation. Through their participation in these meetings the development partners were able to provide more than just financial support for the initiative. In particular, development partners were able to contribute to the project their own expertise and experience in respect of international large scale assessments, capacity building in learning assessment and development project management. The project governance structures that are described later in this report were conducive to incorporating a wide range of development partners in the project and inducting new partners when these came forward in the later years of the project.

---

[3] See the abbreviation and acronym section for the full names of these programmes.

2.1.1. International costs for PISA-D

PISA-D international costs include development of the technical design of the project, its framework and instruments, survey operations, sampling services, translation of project materials, capacity-building activities, data analysis and reporting. PISA-D consists of three technical strands managed by the OECD and overseen by the IAG, and the PISA-D TAG.

Figure 2.2. PISA for Development technical work strands
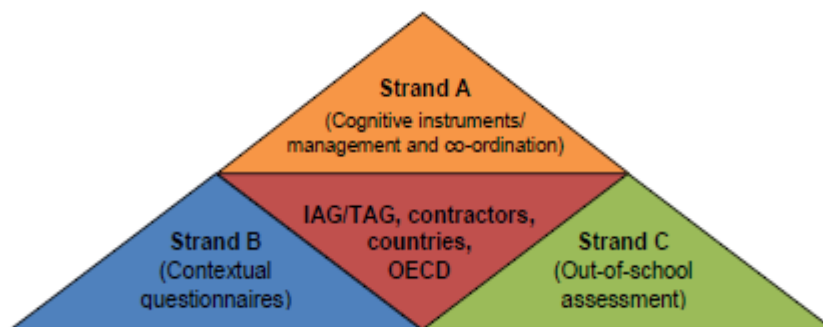
2.1.2. In-country costs for PISA-D

Just like PISA countries, each PISA-D country established a National Centre to implement the assessment in accordance with the PISA technical standards. As already noted, these National Centres were government institutions and are also responsible for most other learning assessments being carried out in the countries. The National Centre is staffed with a team to carry out the necessary tasks, such as provide feedback on the frameworks, conduct sampling activities, translate/adapt survey instruments, collect and manage data, code students' responses, and analyse and report data. The National Centre also had to budget to send staff to attend PISA-D international meetings and trainings.

The in-country budgets are described in a project implementation plan prepared by each country with the support of the OECD. The plans were managed by the countries themselves and, in some cases, their development partners. For example, in addition to supporting PISA-D's international costs, the United Kingdom also supported Zambia's participation in PISA-D; France and the World Bank supported Senegal's participation; and Korea and the Global Partnership for Education supported Cambodia's participation. Some development partners also established extended engagement agreements with PISA-D countries for technical assistance and institutional capacity building.

2.1.3. Funding PISA-D

Project funding for the international costs was made up of a combination of contributions from the PISA-D countries, development partners and private foundations and companies.

Germany, Ireland, Korea, Norway, Porticus Foundation and the United Kingdom contributed funds directly to the OECD to help cover international project costs. The World Bank covered specific costs, including financing three PISA-D background reports and the three-stage capacity-building process to prepare countries for PISA-D implementation: completion of a capacity needs analysis, a capacity-building plan and a

project implementation plan for each country. The Inter-American Development Bank funded Spanish-language capacity-building workshops for the participating Latin American countries as complementary to the workshops conducted in English by the PISA-D contractors. In addition to making a financial contribution, development partners also contributed to the technical discussions regarding the design of the project.

PISA-D also received generous support from the Microsoft Corporation, Positivo and Sunny Varkey in the form of donated tablet computers for conducting the PISA-D out-of-school assessment. Microsoft and Positivo also provided technical support to the contractors to ensure proper maintenance and use of the tablets.

2.2. PISA-D technical meetings and workshops

2.2.1 Initial technical meeting

An initial technical meeting about the PISA-D project was held in June 2013 at the OECD headquarters in Paris, France. The meeting was chaired by the OECD and was attended by representatives from countries committed to or considering participation in PISA-D (Cambodia, Ecuador, Guatemala, Mongolia, Punjab-Pakistan, Senegal, Sri Lanka, Zambia), development partners (France [*Agence française de développement* / French Development Agency], Germany [*Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung* / Federal Ministry for Cooperation and Development and *Deutsche Gesellschaft für Internationale Zusammenarbeit* / German Corporation for International Cooperation], Inter-American Development Bank, Korea, Norway [*Norad er direktoratet for utviklingssamarbeid* / Norwegian Agency for Development Cooperation], United Kingdom [Department for International Development], World Bank), international agencies (UNESCO, UNICEF, Education For All Global Monitoring Report, Conference of the Ministers of Education of French-speaking countries [CONFEMEN] and Programme for the Analysis of Education Systems [PASEC]), and independent experts and technicians.

The objective of the meeting was to make substantial progress towards:

- shared understanding among participants regarding the background to the initiative, its expected results, main challenges, general timeline, the implications of participation for countries and the next steps for partnerships and implementation

- a general agreement regarding the main technical challenges to be addressed by the project

- a framework for the working methods and focus of the technical partnerships with assessment organisations, UN bodies and other agencies

- proposals for membership of the Steering Group, the Technical Advisory Group, and for the first set of technical papers to be commissioned.

Meeting participants concluded the meeting by agreeing on the following steps to develop and implement the PISA-D project:

- OECD to fully develop PISA-D project Terms of Reference (ToR), seek input from partners, and to commission the first set of expert papers; to continue with bilateral dialogues with interested countries to outline country-specific parameters for their participation, including but not limited to capacity-building and funding contributions; to revise the draft participation agreement based on comments

received and to send to interested countries for review and potential approval based on bilateral discussions; to organise a meeting focusing on capacity-building and knowledge-transfer aspects of PISA-D (e.g. Webinar) as part of output 4 of the project; to confirm with other agencies arrangements for technical partnerships and collaboration; to collate information from countries already participating in PISA regarding challenges, opportunities, and capacity-building experiences, specifically regarding student assessment and use of results; to develop a detailed engagement and communication strategy as part of output 5 of the project; to draft ToR for international contractors with input and guidance from participating countries, Steering Group and TAG

- interested countries to work with development partners and the OECD to identify in-country costs associated with the activities that will be conducted by the governments of participating countries and those areas that will need contractors to assume some of the activities; and countries formally participate in PISA-D based on participation agreement

- participating countries and OECD conduct a review of the quality of relevant educational data in each country and the data-collection capacity related to student assessment to inform the capacity-building work programme for each country and the ToRs for International Contractors; and establish agreements with development partners regarding contributions and support (e.g. international costs, in-country costs and activities)

- participating countries nominate their respective National Centres and NPMs and engage with OECD regarding country-specific project planning

- project implementation plans for each confirmed country to be developed in partnership between each participating country, donor partners and the OECD that will include aspects such as a capacity-building work programme, a strategy for how the student assessment results and participation in PISA-D can inform policy discussions in the countries, and priorities for peer-to-peer exchanges and peer learning among countries.

2.2.2 Technical workshops

Three technical workshops – one for each strand of work – were held early in the project to discuss the technical issues to be addressed by the PISA-D project: Strand A – cognitive instruments, Strand B – contextual questionnaires and Strand C – out-of-school assessment. The workshops aimed to bring together experts to reach an agreement regarding the main technical challenges to be addressed and the most technically sound and viable options for addressing these. The workshops also served to define the key components and general structure of the Terms of Reference for the International Contractor(s) to be commissioned for the development and implementation work with participating countries.

The workshops were attended by the OECD, representatives from countries committed to or considering participation in PISA for Development, project development partners, international agencies, and independent experts and technicians

Cognitive instruments

The workshop on PISA-D cognitive instruments took place from 8-9 April 2014 in Washington, D.C., United States. The main resource for the workshop was an expert paper OECD commissioned setting out the technical issues in respect of enhancing the descriptive

power of PISA's cognitive instruments in reading, mathematics and science, particularly with regards to students performing below baseline proficiency levels in PISA, to make these more relevant to contexts in LMIC. The paper, "PISA For Development Technical Strand A: Enhancement of PISA Cognitive Instruments" (OECD Education Working Papers, No. 126) was drafted by Ray Adams and John Cresswell of the Australian Council for Education Research (ACER) and can be found at this link. The workshop covered several topics, including test-targeting, potential "floor effects", PISA framework coverage, item-response types, descriptive power and international comparability to PISA scales, as well as how PISA-D could learn from and build on other relevant assessments.

Contextual questionnaires

The workshop on PISA-D contextual questionnaires took place from 10-11 April 2014 in Washington, D.C., United States. The main resource for the workshop was an expert paper OECD commissioned setting out the technical issues in respect of enhancing PISA's contextual questionnaires, to make these more relevant to LMIC. The paper, "Towards the development of contextual questionnaires for the PISA for development study" (OECD Education Working Papers, No. 118) was drafted by Doug Willms and Lucia Tramonte of the University of New Brunswick and can be found at this link. The workshop covered several topics, including the PISA policy themes relevant to LMIC (selecting and grouping students; resources invested in education; school governance, assessments and accountability; and the quality of learning environments) and the themes to focus on for enhancing the contextual questionnaires for PISA-D (early learning opportunities, family and community support, learning time, socioeconomic status, and school resources) , as well as how PISA-D could learn from and build on other relevant assessments.

Out-of-school assessment

The workshop on designing the main components of PISA's out-of-school survey took place from 1-2 October 2014 in Montreal, QC, Canada. The main resource for the workshop was an expert paper OECD commissioned setting out the technical issues in respect of establishing methods and approaches to include out-of-school youth in the assessment. The paper, "PISA For Development Technical Strand A: Enhancement of PISA Cognitive Instruments" (OECD Education Working Papers, No. 120) was drafted by independent expert Roy Car-Hill and can be found at this link. The workshop covered several topics, including the definition and identification of out-of-school 15-year-olds, sampling and surveying approaches, country-specific expectations, survey administration and response rates, and the cognitive and contextual assessments.

2.2.3 TAG and Expert Group meetings

The PISA-D TAG met in September and November 2015, June 2016, May and October 2017, and April and August 2018. The meetings were attended by the OECD, the PISA-D contractors, and the TAG members: experts in appropriate fields (drawn mainly from the existing TAG for main PISA) and those individuals who (i) have a leading operational role in PISA-D; and (ii) have expertise in technical areas that are not available from the membership of the existing TAG for main PISA, such as surveys of out-of-school youth in LMIC countries. Meetings topics included policy needs that PISA-D could effectively meet, overall programme design, methodological approaches for data collection and

surveys to include out-of-school 15-year-olds in the assessment, implementation of the programme, and finalisation and use of the PISA-D database.

The PISA-D Subject Matter Expert Group (SMEG) met in July 2015 in Mexico City, Mexico to: review the approach and methodology in extending the PISA reading, mathematics and science frameworks for PISA-D; review item pools to understand characteristics of PISA items at the lower end of the proficiency scale; gain understanding of 15-year-old students' abilities in the potential PISA-D countries; discuss framework validation and the operationalisation of the frameworks in assessment tasks; and reach consensus on the approach to the PISA-D reading, mathematics and science frameworks. The SMEG consists of representatives from participating countries to ensure each of the frameworks fits the purpose of assessing 15-year-olds in middle- and LMIC. In addition to the face-to-face meeting in 2015, the SMEG met virtually over the course of the PISA-D project and worked closely with the PISA-D contractors to steer the direction of the frameworks.

The PISA-D Questionnaire Expert Group (QEG) met in August 2015 in Stellenbosch, South Africa to discuss: the PISA-D contextual framework; the questionnaire item map; measures of Prosperity Outcomes (educational attainment, health and well-being, and student engagement); measurement of socioeconomic status and poverty; measures of Foundations for Success (inclusive environments, learning time, quality instruction, family and community support, and resources); and assembling the student, teacher and parent questionnaires. The QEG consists of representatives from participating countries to ensure each of the questionnaire framework fits the purpose of assessing 15-year-olds in LMIC. In addition to the face-to-face meeting in 2015, the QEG met virtually over the course of the PISA-D project and worked closely with the PISA-D contractors to steer the direction of the framework.

2.3. PISA-D IAG and NPM meetings

The PISA-D IAG met annually over the course of the project:

- 1st meeting – May 2014 – OECD Headquarters, Paris, France: Established the Terms of the Reference (TOR) for the IAG, including annual report content and election of two Co-Chairs; reported on PISA-D project progress; discussed and agreed on the main technical issues to be addressed and the best options to tackle the issues; discussed and agreed on the TOR for the International Contractors to be commissioned by the OECD for project implementation; reviewed the capacity needs analysis conducted with participating countries and the design of the capacity building programmes; and reviewed the proposal for an independent review of the project to be conducted in 2016/17.

- 2nd meeting – March 2015 – OECD Headquarters, Paris, France: Adopted IAG TOR; reviewed annual report on project activities; agreed on the PISA-D TAG TOR; approved the International Contractors' Statements of Work and integrated timelines for Strands A and B; approved capacity building preparation of participating countries; discussed commissioned report drafts on the experiences of middle-income countries in PISA, the status of system-level data in PISA-D countries, and the review of large-scale assessments in education; adopted the PISA-D engagement and communication strategy; agreed that the independent review of the project would be conducted in 2017/18; and initiated peer-learning.

- 3rd meeting – March 2016 – Asunción, Paraguay: Reviewed annual report on project activities; presentation by International Contractors of detailed progress reports for Strands A, B and C; PISA-D countries reported on capacity building activities as well as stakeholder engagement and communication activities; peer learning with Kosovo on communicating with stakeholders, preparing a national report and disseminating results; discussed the PISA-D analysis and reporting plan; discussed the status of system-level data in PISA-D countries; agreed that the independent review of the project would be conducted in 2019; discussed scaling up of PISA-D as described in the PISA Development Strategy.

- 4th meeting – May 2017 – Siem Reap, Cambodia: Reviewed annual report on project activities; presentation by International Contractors of detailed progress reports for Strands A, B and C; reviewed the outcomes from the field trial for Strands A and B; reviewed the PISA-D Assessment and Analytical Framework as well as the analysis and reporting plan; PISA-D countries reported on capacity building activities as well as stakeholder engagement and communication activities; peer learning with Brazil and Kosovo on communicating with stakeholders, preparing a national report and disseminating results; agreed that the independent review of the project would be conducted in 2018/19; discussed regional assessment experiences with NEQMAP and SEA-PLM.

- 5th meeting – May 2018 – Saly, Senegal: Reviewed annual report on project activities as well as progress reports for Strands A, B and C; reviewed the analysis and reporting plan; reviewed the outcomes from the field trial for Strand C; PISA-D countries reported on capacity building activities as well as stakeholder engagement and communication activities; peer learning with France and Luxembourg on communicating with stakeholders, preparing a national report and disseminating results; agreed that the independent review of the project would be conducted in 2019; discussed plans for an end-of-project international seminar in October/November 2019; and discussed regional assessment experiences with regards to analysing data, reporting and evidence-based decision-making.

Nine PISA-D NPM meetings were held over the course of the project to build country capacity in the following areas:

1. September 2015 – Quito, Ecuador: PISA cognitive and contextual frameworks, characteristics of the available item pools

2. January 2016 – Rockville, Maryland, United States: Adaptation, translation and verification of survey materials and sampling

3. April 2016 – Asunción, Paraguay: Student sampling and field trial survey operations

4. July 2016 – Livingston, Zambia: Scoring and coder training and data management for the field trial

5. November 2016 – Madrid, Spain: Quality control sample selection forms, and quality assurance procedures, data management (software, codebook, etc.)

6. May 2017 – Siem Reap, Cambodia: Analysis and interpretation of field trial results and preparation for Main Survey

7. July 2017 – Princeton, New Jersey, United States: Student sampling and Main Survey operations

8. May 2018 – Saly, Senegal: Data processing, results, scaling methodology and preparation for analysis

9. July 2018 – Antigua, Guatemala: Analysis and interpretation of Main Survey results, reporting and dissemination of results

These NPM meetings were scheduled to support project implementation in accordance with the key phases outlined at the beginning of this chapter, namely:

*   Design, planning and co-ordination (2014-16);

*   Technical development (2015-16);

*   Field trials and in-country data collection (2016-18);

*   Analysis and report writing (2018-19); and

*   Report production, dissemination and post-pilot governance (2018-20).

In addition, the NPM meetings were supplemented by visits to the participating countries by the contractors for capacity building and support, peer-to-peer learning activities and support from the OECD Secretariat. These capacity building activities are described in more detail at Chapter 6 of this report.

There were two other key meetings that were implemented as part of the project and these were as follows:

*   The PISA-D International Seminar took place on 25th September 2019 at the Westminster Central Hall, London, United Kingdom. The Seminar was held in conjunction with the 48th Meeting of the PISA Governing Board (PGB). Representatives of the following attended the meeting either in person or remotely: countries participating in PISA and PISA-D; development partners; institutional partners; international contractors; civil society; academic institutions; research institutions; and independent experts. The Seminar discussed the extent to which the five outputs of the PISA-D project have been achieved and identified the key lessons that have been learned, particularly in respect of making PISA more accessible to and relevant for a wider range of countries and putting Low-and-Middle-Income-Countries on the PISA scale. The seminar also considered the wider implications of the project and its outcomes for education systems development and learning assessment as an integral element of effective systems in the context of SDG 4.

*   The PISA-D virtual International Workshop took place on 3rd December 2020. Representatives of the following attended the meeting remotely: countries participating in PISA and PISA-D; development partners; institutional partners; international contractors; civil society; academic institutions; research institutions; and independent experts. The Workshop discussed the results of Strand C of the PISA-D project (assessment of the out-of-school youth) and considered the options for scaling up the pilot, particularly the option of integrating the assessment with household surveys.

2.4. Project implementation was successful

Strands A and B of the project were implemented successfully and in accordance with the plan and timeline agreed at the outset of activities and in accordance with PISA's technical standards. Strand C of the project was designed to commence six months after the commencement of Strands A and B as it was expected that the managerial burdens for the countries and the contractors would be too great if they were implementing the in-school and out-of-school surveys at exactly the same time.

Strand C of the project was extended by 18 months for three reasons:

- First, a three month extension was agreed in 2017 on account of the longer than expected time it took to deliver and clear the tablet computers through the customs of the participating countries.

- Second, a six month extension was agreed in 2018 when it became clear that the participating countries would not complete their Main Survey data entry and coding activities before February 2019.

- Third, a nine month extension was agreed in 2020 because of the impact of the Covid-19 global pandemic.

The key factors that have contributed to a successful implementation of the project, include:

- An effective project design, including the necessary support for capacity needs analysis, capacity building planning and support for implementation for the participating countries (see Chapter 6 for more details).

- Strong project governance and management arrangements led by the OECD, in particular the IAG, that facilitated effectively the contributions of a large number of development partners, participating countries, institutional partners, contractors and external experts.

- Strong ownership of the PISA-D process by the national governments of participating countries with the necessary institutional support provided for implementation.

- Strong commitment and performance of the National Centres, National Project Managers and national teams.

- Excellent project management and technical inputs from the international contractors.

- The TAG and Expert Groups functioned effectively.

- The institutional and technical partnerships worked well and their expertise was made available in thematic areas (e.g. with UNICEF and UIS for out-of-school youth).

- Development partners delivered their support (international and in-country) and technical guidance on time and helped the project to move forward.

- PISA participating countries were generous and valuable contributors to peer-to-peer learning.

3. PISA cognitive instruments: Descriptive power of reading, mathematics and science enhanced

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

PISA has six proficiency levels, with Level 6 the highest and Level 1 and below the lowest. Level 2 is a particularly important threshold, as this marks the baseline level of proficiency at which students begin to demonstrate the competencies that will enable them to participate effectively and productively in life as continuing students, workers and citizens.

The OECD analyses reported in the PISA 2012 volumes show that the overall performance of 15-year-old students in all of the LMIC participating in PISA 2012, except Viet Nam, was lower than that of students in OECD countries, and varies widely. Performance is also concentrated at the lower levels of the PISA proficiency scales. For example, for more than half the LMIC participating in PISA 2012 scores are concentrated at Level 1 or below on the mathematics proficiency scale.

In previous PISA cycles, the distribution of items on the proficiency scales focused more on Levels 3 and 4. Since a major aim of PISA-D is to provide a wider and more fine-grained picture of what 15-year-olds in LMIC know and can do at these lower levels of performance, it was necessary to augment the distribution of items in the assessment to allow for more accurate measurement of proficiency at Levels 2 and below. To achieve this goal, PISA-D included more cognitive items at Level 2 and below. In PISA 2018, 22% of the items were at Level 2 and below compared with 65% of the items in PISA-D. This broader set of information about performance at and below the baseline level of proficiency will be directly relevant to curriculum planners, teacher educators and other education professionals in LMIC in designing improvement plans and policies.

While the PISA-D test design and items target the lower levels of performance, the assessment links to the whole of the PISA frameworks for comparability. This link is established through the inclusion of a larger proportion of trend items from the PISA 2015 paper-based assessment. PISA-D further differentiates the lower end of the proficiency scale for each subject tested (reading, mathematics and science) while maintaining measurement of the higher levels.

The development of the PISA-D Strand C cognitive assessment design is discussed later in this report. However, it is important to note briefly that the Strand C test focused on Reading and Mathematics only, with equal weights for each domain (i.e. no major/minor domain distinction as is made in PISA). More importantly, the Strand C test was computer-based and linked to PISA, using a subset of the items chosen for PISA-D Strand A with the distribution of items selected for PISA-D Strand C focused on the lower end of the difficulty scale.

3.1. The PISA-D school-based assessment design

PISA D assessment instruments were developed with the goal of providing reliable, valid, and comparable information from students in a wide range of LMIC while ensuring that results are linked to the main PISA assessment. This design relied on the administration of paper-based assessment materials for 15-year-old students in grades 7 and above, as well as context (background) questionnaires for school administrators, teachers, and the person(s) most knowledgeable about the student (i.e. parent or guardian).

The domain coverage specified in the assessment design was intended to extend the range of information that PISA would provide to policy makers concerning the distribution of skills in their student populations. In summary, PISA D was designed to provide participating countries with the following information:

- population distributions in Reading, Mathematics, and Science that reflect the PISA-D frameworks

- links to the proficiency scales reflected in the PISA 2015 paper-based assessment.

The assessment design includes four clusters of items from each of the domains of reading, mathematics and science. In order to meet the goals and domain coverage, each cluster was assembled from a combination of intact units of items from PISA 2015 along with items selected from existing surveys. The items within each cluster represented a range of key framework aspects, item types, and item difficulties.

There are 12 different test booklets in the Main Survey, each containing items from two of the three core PISA domains. Each booklet allocated to students comprises four 30-minute clusters of test material. In total, students were expected to spend 120 minutes responding to a set of items from one of the 12 booklets. This timing is consistent with the timing for the main PISA assessment.

Each test booklet is completed by a sufficient number of students to make appropriate estimates of the achievement levels on all items by students in each country and in relevant subgroups within a country (such as boys and girls, and students from different social and economic contexts). Comparability with the PISA 2015 paper-based assessment is assured through linkage with the trend items. In addition, each student answers a 35-minute background questionnaire, which gathers contextual information that is analysed with the test results to provide a broader picture of student performance.

3.1.1. An overview of what is assessed in each domain

Box 3.1 presents definitions of the three domains assessed in PISA-D, which are taken from the assessment frameworks used in PISA 2015 (OECD, 2016). The definitions all emphasise functional knowledge and skills that allow one to participate fully in society. Such participation requires more than just being able to carry out tasks imposed externally by, for example, an employer; it also means being able to participate in decision making. The more complex tasks in PISA-D require students to reflect on and evaluate material, not just to answer questions that have one correct answer.

Box 3.1. Definitions of the domains

Reading literacy: An individual's capacity to understand, use, reflect on and engage with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.

Mathematical literacy: An individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens.

Scientific literacy: The ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

3.2. Measuring reading literacy in PISA-D

PISA defines reading literacy as: understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, develop one's knowledge and potential, and participate in society. This definition acknowledges the diversity and complexity of the processes involved in daily reading activities. PISA also establishes a baseline level – proficiency Level 2, on a scale with 6 as the highest level and 1b the lowest – at which readers begin to demonstrate the competencies that will enable them to participate effectively and productively in life as continuing students, workers and citizens.

As PISA was originally developed for middle- and high-income countries, proficiency below this baseline was less clearly defined and there are fewer tasks included in the item pool that represented these levels of proficiency.

The extensions made to the PISA frameworks for PISA-D are an attempt to gain more information about students at the bottom of the performance distribution, particularly for those students performing at or below Level 1b. PISA-D attempts to refine the description of reading proficiency at the lowest range of the reading scale in PISA 2009 (below Level 1b) by making a new distinction of a Level 1c. Tasks at Levels 1c require the reader to demonstrate understanding of literal meaning of words and phrases within very short, syntactically simple passages. The texts represented at Level 1c are short and include little, if any, competing information. At Level 1b, texts usually require the reader to understand the meaning of sentences within a short, simple passage with limited competing information.

The PISA reading assessment is built on three main task characteristics:

- Processes (aspects) – the cognitive strategies, approaches or purposes that readers use to negotiate their way into, around and among texts

- Text – the range of material that is read

- Situation – the range of broad contexts or purposes in which reading takes place.

Five main reading processes were defined in the reading literacy framework: retrieving information; forming a broad understanding; developing an interpretation; reflecting on and evaluating the content of a text; and reflecting on and evaluating the form of a text.

While the texts and situations in the PISA-D reading assessment are the same as those used in PISA 2015, the reading tasks in PISA-D have a greater focus on processes that highlight foundational reading strategies or skills that are required to perform successfully at the lowest levels of the PISA reading proficiency scale.

With a clearer understanding of the reading skills of low-performing 15-year-olds, policy makers in LMIC will be able to design more effective and targeted policies to help students learn to read better, help teachers teach reading better, and help school systems promote reading as a fundamental skill.

3.2.1 Strategy to provide better measurement of basic literacy levels

Two strategies were used to expand the item pool in the reading assessment to better measure lower levels of reading proficiencies as defined in the extended framework. First, additional item types were included to assess basic sentence and passage comprehension. Sentence processing tasks assess the ability to understand the basic meaning of simple sentences of varying lengths and syntactic complexity. In the PISA-D reading components section of the reading assessment, the construct is instantiated in a sensibility judgment task. Its purpose is to measure the extent to which students can comprehend sentences of increasing lengths. In these tasks, students see a set of sentences and decide if they make sense ("yes") or do not make sense ("no") with respect to general knowledge about the real world, or the internal logic of the sentence itself.

The basic passage comprehension tasks assess the ability to understand the literal meaning of short simple sentences in a text. This aspect of the construct has been instantiated through a set of tasks in which the participant must complete a sentence by selecting a word from among three options. Its purpose is to measure the extent to which students understand the literal and inferential meaning of connected text. Sample item 1 shows an example to illustrate the passage comprehension task.

Other tasks to assess the lower end of the proficiency scale involved adapting existing PISA tasks to assess low-level access and retrieve processes. Sample item 2 illustrates this.

3.2.2 Testing reading literacy among the out-of-school population

The extended PISA-D reading framework is appropriate for 15-year-old students whether in or out of school. The units and items are not directly based in the school context, and thus there is no particular requirement or change needed in the units that are categorised as relevant for educational activities, since educational activities also occur out of school. Therefore, the distribution and selection of units and items can be the same for PISA-D in-school and out-of-school populations. The out-of-school component is assessed on a tablet computer, using units that are a subset of those included in the school-based assessment.

3.2.3 Examples of items for addressing the extended PISA-D reading framework

Sample item 1 – Passage comprehension

In items assessing passage comprehension, respondents are asked to read a passage in which they are required at certain points to select the word that makes sense from the two alternatives provided.

To the editor: Yesterday, it was announced that the cost of riding the bus will increase. The price will go up by twenty percent starting next <u>wife / month</u>. As someone who rides the bus every day, I am upset by this <u>foot / increase</u>. I understand that the cost of <u>gasoline / student</u> has risen. I also understand that riders have to pay a fair <u>price / snake</u> for bus service. I am willing to pay a little more because I rely on the bus to get to <u>object / work</u>. But an <u>increase / uncle</u> of twenty percent is too much.

This increase is especially difficult to accept when you see the city's plans to build a new sports stadium. The government will spend millions on this project even though we already have a <u>science / stadium</u>. If we delay the stadium, some of that money can be used to offset the increase in bus <u>fares / views</u>. Then, in a few years, we can decide if we really do need a new sports <u>cloth / arena</u>. Please let the city council know you care about this issue by attending the next public <u>meeting / frames</u>.

Sample item 1 assesses passage comprehension and reflects a typical Level 1c task. In PISA-D, the passage comprehension paragraphs have been modified to have three options instead of two.

Sample item 2 – Supermarket notice

Peanut Allergy Alert

Lemon Cream Biscuits

Date of alert: 04 February

Manufacturer's Name: Fine Foods Ltd

Product Information: 125g Lemon Cream Biscuits (Best before 18 June and Best before 01 July)

Details: Some biscuits in these batches may contain pieces of peanut, which are not included in the ingredient list. People with an allergy to peanuts should not eat these biscuits.

Consumer action: If you have bought these biscuits you may return the product to the place of purchase for a full refund. Or call 1800 034 241 for further information.

QUESTION

What is the name of the company that made the biscuits?

……………………………………………………………………….

Sample item 2 is a released PISA item that assesses basic access and retrieve process that is typical of a Level 1b task, as the passage is short and the task requires a small inference to understand that "manufacturer" and "company" are synonymous and locate a piece of explicitly stated information with limited competing information to identify the company's name.

3.3. Measuring mathematics literacy in PISA-D

PISA defines mathematical literacy as: formulating, employing and interpreting mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assesses the capacity of individuals to recognise the role that mathematics plays in the world and to make the well-founded judgements and decisions needed to be constructive, engaged and reflective citizens. As with reading, proficiency Level 2 in mathematics, is the level at which individuals begin to demonstrate the competencies that will enable them to participate effectively and productively in life as students, workers and citizens.

The conceptual framework for the mathematics assessment in PISA-D expands the descriptions of the proficiencies and processes to better differentiate the skills necessary at the lower end of the mathematics proficiency scale. These enhancements to the PISA mathematics framework in PISA-D focus on three aspects of the assessment in particular:

1. *Proficiencies*: PISA-D differentiates performance at the lowest level by breaking down Level 1 into three sub-levels: 1a, 1b and 1c.

2. *Proce*sses: To better describe students' attempts to apply mathematical processes, PISA-D extends the descriptions of the processes to include:

   ● "Selecting an appropriate model from a list" to the PISA mathematical process "formulate situations mathematically"

   ● "Performing a simple calculation", "drawing a simple conclusion" and "selecting an appropriate strategy from a list" to the PISA mathematical process "employ mathematical concepts, facts, procedures and reasoning"

   ● "Evaluating a mathematical outcome in terms of the context" to the PISA mathematical process "interpret, apply and evaluate mathematical outcomes".

3. *Skills*: PISA-D adds and assesses 15-year-olds' ability to "select a model appropriate to the context of real-world problems", "select a representation appropriate to the context", "select an appropriate justification" and "implement a given strategy" as some of the skills required to solve the problems presented in PISA-D.

Tasks at Levels 1b and 1c require responses using simple calculations in a single step or operation. The contexts of the tasks are easy to understand, and all relevant information is

provided in a simple, familiar format (e.g. short paragraph or table) with limited information.

With a clearer understanding of the mathematical skills of low-performing 15-year-olds, policy makers in LMIC will be able to design more effective and targeted policies to help students perform basic mathematical operations with greater ease and understanding, help teachers teach mathematics better, and help school systems promote mathematics as a fundamental skill.

3.3.1 Strategy to provide better measurement of basic mathematical literacy levels

The key strategy for measuring the lower level of the proficiency scale is to include more items at and below Level 2 with familiar contexts and that use more straightforward language and simply formulated concepts. In order to gain useful information for the new Levels 1b and 1c it is vital that context and language do not interfere with the mathematics being assessed. To this end, the context and language is carefully considered.

The context for both 1b and 1c should be situations that students encounter on a daily basis. Examples of these contexts may include money, temperature, food, time, date, weight, size and distance. All items should be concrete and not abstract. The focus of the item should be mathematical only. The understanding of the context should not interfere with the performance of the item.

Equally important is to have all items formulated in the simplest possible terms. Sentences should be short and direct. Compound sentences, compound nouns and conditional sentences should be avoided. Vocabulary used in the items must be carefully examined to ensure that students will have a clear understanding of what is being required. In addition, special care will be given to ensure that no extra difficulty is added due to a heavy text load or by a context that is unfamiliar to students based on their cultural background.

Items designed for Level 1c should only ask for a single step or operation. However, it is important to note that a single step or operation is not limited to an arithmetical step. This step might be demonstrated by making a selection or identifying some information. All parts of the modelling cycle can and should be used to measure the mathematical abilities of students at Levels 1b and 1c.

3.3.2 Testing mathematical literacy among the out-of-school population

For the out-of-school population, item selection focused on the scale at or below Level 2 with an emphasis on the lower end of the scale in terms of item distribution. The selection process was similar to that used for the in-school population: coverage of all processes was maintained and contexts of the items were reviewed to ensure appropriateness for what individuals would encounter in an out-of-school context.

3.3.3 Examples of items for addressing the extended PISA-D mathematics framework

Sample item 1 – Walking
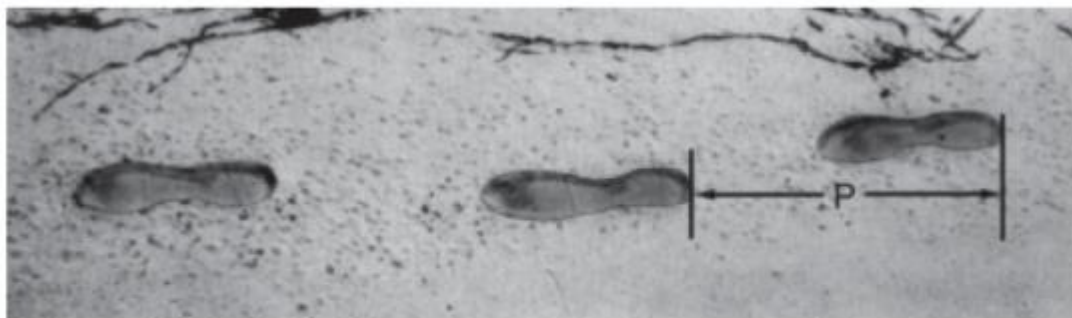
The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula $\dfrac{n}{P} = 140$ gives an approximate relationship between n and P where

    n = number of steps per minute, and
    P = pacelength in metres.

QUESTION

Heiko has a pacelength that is 0.5 metres. Using this formula, how many steps per minute, n, does Heiko take each minute?

………………………………………………

The item represents the type of task that is typical at Level 1a. The context is familiar and the question is clearly defined. A simple substitution is required to be carried out in a single operational step.

Sample item 2 – Exchange rate

Mei-Ling found out that the exchange rate between Singapore dollars and South African rand was

<div align="center">1 SGD = 4.2 ZAR</div>

Mei-Ling changed 3 000 Singapore dollars into South African rand at this exchange rate. Choose a correct method from those listed. Then calculate *n*, the amount of South African rand Mei-Ling received after the exchange.

$$\frac{1}{4.2} = \frac{n}{3000} \qquad \frac{1}{3000} = \frac{4.2}{n} \qquad 4.2n = 3000 \qquad n = 3000(4.2)$$

For this item, the student is given four methods to solve for *n*. Two of these methods will result in a correct value for *n*. The expectation is that a student will be able to select one of the correct methods and then solve for the value of *n*. This addresses the added process, "selecting an appropriate model from a list." If a student is able to choose one of the correct

methods, the requirements for proficiency 1b are met. If a student is also able to solve for *n* correctly, the requirements for proficiency 1a are met.

3.4. Measuring science literacy in PISA-D

PISA defines scientific literacy as the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. PISA's definition includes being able to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically. It emphasises the importance of being able to apply scientific knowledge in the context of real-life situations. As with reading and mathematics, proficiency Level 2 in science, on a scale with 6 as the highest level and 1b the lowest, is the level at which individuals begin to demonstrate the competencies that will enable them to participate effectively and productively in life as students, workers and citizens.

The conceptual framework for the science assessment in PISA-D extends and broadens the measurement at the lower end of the performance spectrum by including more items at and below Level 2 difficulty. For the items at the lowest possible proficiency level, the questions are formulated in the simplest possible language to reduce the cognitive demands on students and whenever possible draw on phenomena that are familiar to students' everyday lives or on ideas that are pervasive in contemporary culture. These enhancements to the PISA science framework in PISA-D focus on three aspects of science literacy in particular:

1. *Proficiencies* – PISA-D establishes Level 1c as the new lowest level on the science proficiency scale, and attempted to include test items to measure performance on tasks at this level.

2. *Competencies* – PISA-D provides a detailed description of each of the three PISA competencies for proficiency Levels 1a, 1b and 1c. For example, for the competence "Explain phenomena scientifically":

   ● Level 1c items require recognising explanations for a limited range of the most simple natural and technological phenomena demonstrating the ability to recall appropriate scientific knowledge.

   ● Level 1b items require recognising explanations for a range of simple or familiar natural and technological phenomena demonstrating the ability to identify an explanatory model or representation, and recognise the potential implications of scientific knowledge for society and individuals.

   ● Level 1a items require recognising explanations for a range of simple or familiar natural and technological phenomena demonstrating the ability to make appropriate predictions, recognise an appropriate explanatory hypothesis and recognise simple causal or correlational relationships.

3. *Skills* – PISA-D identifies the skills necessary to perform at the lowest proficiency level as: the student must be able to read and comprehend simple sentences, use numeracy and basic computation, understand the basic components of tables and graphs, apply the basic procedures of scientific enquiry and interpret simple data sets.

With a clearer understanding of the scientific skills of low-performing 15-year-olds, policy makers in LMIC will be able to design more effective and targeted policies to help students

improve their ability to engage with science-related issues, help teachers teach science better, and help school systems promote science as a fundamental skill.

3.4.1 Strategy to provide better measurement of basic science literacy levels

A strategy for measuring the lower level of the science proficiency scale is to expand the item pool at and below Level 2 with familiar contexts and that use more straightforward language and simply formulated concepts, similar to the strategy for mathematics. For science, this strategy involves identification of more fundamental scientific phenomena in relatable contexts.

Items at the newly defined Level 1c should be familiar to students' everyday lives or draw on ideas that permeate contemporary culture. All items should, whenever possible, attempt to draw on macroscopic phenomena that students may have experienced or observed or learnt in the curriculum. Equally important is to have all items formulated in the simplest possible language. Sentences should be short and direct. Lengthy sentences, compound nouns and complex phrasing should be avoided. Vocabulary used in the items must be carefully examined to avoid the use of academic language and, wherever possible, simplify the scientific language. In addition, the cognitive processing should only require one-step reasoning and use simple data or descriptions.

Students with proficiency scores at Level 1c require the foundational skills to:

- read and comprehend simple sentences

- use numeracy and basic computation

- understand the basic components of tables and graphs

- apply the basic procedures of scientific enquiry

- interpret simple data sets.

3.4.2 Testing scientific literacy in the out-of-school population

The scientific literacy domain is not included in the out-of-school PISA-D assessment due to practical reasons related with the instrument and administration constraints. The total test allows a maximum of 50 minutes, which is not enough time to include an assessment of three domains, so it became necessary to choose only two. In deciding on the domains to include, it was taken into account that reading and mathematics literacy are considered as foundational skills and necessary for the development of scientific literacy skills. In addition, the target population was also considered. Science is the domain with the strongest link to school, so the least appropriate for a group that, by definition, has been exposed to less formal schooling. Thus, it was decided that reading and mathematics were the only domains that should be included in the assessment for out-of-school 14-16 year-olds.

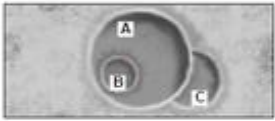### 3.4.3 Examples of items for addressing the extended PISA-D science framework

Sample item 1 – Meteoroids and craters

Question 3A reflects a typical item at Level 1b. This is a basic data interpretation question that requires simple, everyday knowledge that a larger object would cause a larger crater and a smaller one would cause a smaller crater.

Question 3B is an example of a Level 2 item. This question is somewhat more difficult because students must compare the three craters shown in the image to determine when the craters were formed, from oldest to newest, based on the way they overlap in the image – e.g. crater C must have formed first because crater A overlaps C a bit and crater B must be the most recent crater because it is within A.

Sample item 2 - Death of bee colonies

Scientists believe there are many causes why bee colonies die. One possibility is an insecticide that may cause bees to lose their sense of direction outside the hive.

Researchers tested whether this insecticide leads to the death of bee colonies. In a number of hives, they added the insecticide to the food of the bees for three weeks. All of the hives were given the same amount of food but the food had different amounts of insecticide in. Some hives were not given any insecticide.

None of the colonies died immediately. However, by week 14, some of the hives were empty. The following graph records the results:
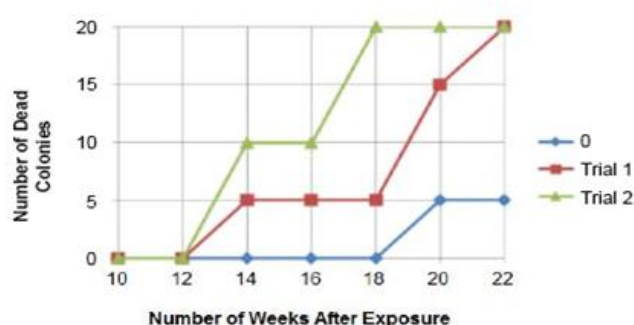
QUESTION

What did the experiment test? Choose one of the responses below:

A. The experiment tested the effect of insecticide on the resistance of bees over time.

B. The experiment tested the effect of varying amounts of insecticide on the number of empty hives found over time.

C. The experiment tested the effect of the death of bee colonies on the resistance of bees to insecticide.

D. The experiment tested the effect of the death of bee colonies on the concentration of the insecticide.

Although the stimulus text is somewhat challenging, the demands of the question are representative of Level 1a items. For this question, students must identify a question being asked in a simple scientific enquiry where only one factor is being varied at a time.

3.5. Findings from the PISA-D field trial of the school-based cognitive test[4]

The PISA-D school-based tests were piloted in seven countries[5]: Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia. The field trial of the school-based assessment instruments took place from August to December 2016. Based on results of the

---

[4] The findings from the PISA-D field trial of the out-of-school assessment are discussed in Chapter 5 of this Report

[5] Bhutan was not included in the pilot from August to December 2016 as the country joined the project late, nor were the country's Field Trial results used to make decisions for the Main Survey.

field trial, the instruments and survey operations were modified as necessary for the Main Survey.

The PISA-D school-based assessment field trial revealed that the instruments work as intended. Findings from the analysis of the PISA-D cognitive test field trial data are organised around three major goals and are summarised in the table below.

Table 3.1. Findings from the field trial of the PISA-D cognitive test

Cognitive test field-trial goals	Findings
Provide information about the data yield and survey operations	All countries met the field-trial sample size requirements. They submitted data to the PISA-D contractors to evaluate data quality and whether the intended goals regarding the functioning of the items were met. Analysis of five countries' data showed good quality and that the cognitive items worked well.
Assess the quality of the items that are either borrowed or adapted from PISA or other international assessments	PISA-D items behaved as expected. A wide range of item difficulty is covered, suggesting an adequate distribution of items: easy (56% of the total), medium (33%) and difficult (11%). No systematic differences in the behaviour of test items were observed across countries, suggesting that the assessment worked in a similar way in all participating countries.
Evaluate field-trial data with item-response theory models to establish reliable, valid and comparable scales.	PISA-D cognitive test field-trial data are comparable with PISA. A strong linkage of the PISA-D scales could be established across countries and to PISA 2015.

The results of the field trial also confirmed that the participating countries carried out their survey operations in accordance with PISA's technical standards. The field trial data were collected by all the countries in a consistent, reliable and valid fashion using test materials that met PISA's high quality standards – these materials had been translated, adapted and verified from the original sources. The countries each submitted a single database to the OECD's international contractors for processing and this data are internationally comparable. All findings from the PISA-D field trial, including lessons learnt, were documented for reference in preparing for the main data collection. While the results of the field trial were positive and its goals were accomplished, each country faced its own challenges in conducting the school-based assessment, such as completing all field-trial tasks on time. The field trial helped countries better plan for the Main Survey and anticipate challenges. The OECD and its contractors provided countries with tailored support to reinforce their capacity and create the conditions in each country that are needed to successfully implement the assessment.

3.6. Findings from the PISA-D Main Survey of the school-based cognitive test[6]

The main survey data collection is subject to a strict adjudication process, particularly for the sampling and translation/adaptation parts of the implementation. The way the sample is selected and the instruments are localised in each country is an integral part of the evidence supporting the construct validity of the intended uses of the assessment. Overall, the PISA-D adjudication review suggests that the PISA-D cognitive test has been implemented in accordance with PISA's technical standards and that the isolated deviations from these standards did not affect the quality of the data products. The quality with which procedures were implemented in PISA-D is in fact comparable to any previous cycle of

---

[6] The findings from the PISA-D Main Survey of the out-of-school assessment are discussed in Chapter 5 of this Report.

PISA. In particular, countries complied with standards related to all aspects of PISA-D implementation with the exception of the following deviations:

- with the exception of one country, countries did not meet the sample size minimum due to poor frame information for the number of PISA-D eligible students, but there is no indication of bias in any of the samples and this did not affect compliance with the remaining sampling standards

- there were some minor issues regarding quality monitoring in one country

- there were some minor issues regarding data submission in three countries

- there were problems with English language capacity in one country.

The minor deviations from the standards summarised above have been documented in the adjudication database and described in more detail in the PISA-D Technical Report (see www.oecd.org/pisa/pisa-for-development/pisafordevelopment2018technicalreport/). These minor deviations were discussed at the PISA-D TAG meeting, and it was noted that these had no impact on the quality of the data which was found fit for reporting in accordance with the goals of PISA-D and, most importantly, for reporting countries' average performance on the PISA scale and for reporting students' proficiency, particularly at Level 2 and below. There were no major issues that required adjudication.

During the analysis phase of the project, analysts took an in-depth look at the results, including the functionality of all test items, and confirmed that the instruments measure what they purport to measure against the *PISA-D Assessment and Analytical Framework* (OECD, 2018[5]). Indeed, the analysis validated the assumptions that shaped the framework. As a result, participating countries can be confident that the assessment results will provide relevant data that can inform decisions concerning national policies.

3.6.1 International comparisons

PISA-D was successful in making the assessment instruments more relevant to LMIC while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA. These variables and factors include student performance (e.g. Figure 3.1), educational attainment, health and well-being, attitudes towards school and learning, the learning environment, learning time, the quality of instruction, family and community support, and resources devoted to education.

Figure 3.1. Students' proficiency in reading, PISA 2015 and PISA-D
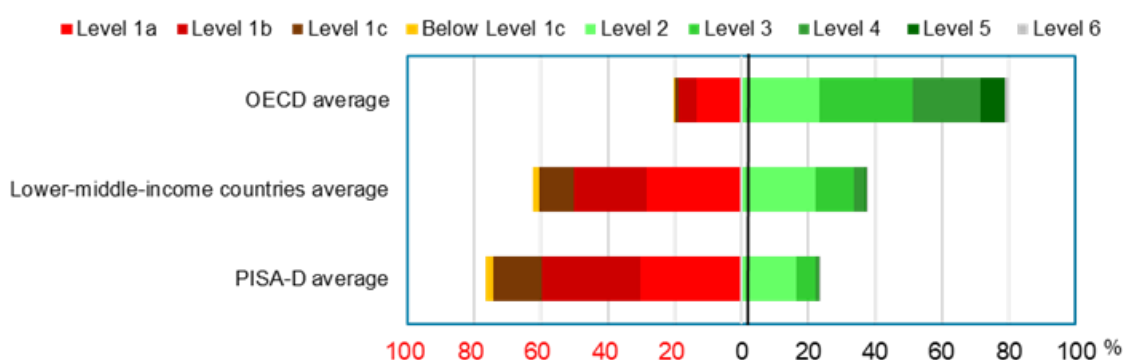
Source: PISA 2015 Database and PISA for Development Database.

3.6.2 Policy insights

PISA-D assessment results provide countries with a solid database that can help them refine policy priorities and set new goals or targets to improve their education systems. The data collected have a lot to say about the allocation of resources and its implications for equity. With reliable data on gaps in access and differences in outcomes between groups of children and young people, countries can determine whether poor and marginalised populations are given equal opportunities to succeed at school and beyond. The challenge for countries is to maintain a focus on these goals or targets, and to track progress towards them by participating in future cycles of PISA and other relevant studies.

3.7 Achievement of the output

This output of the project has been achieved on time and within the scope of the project budget. The project has been successful in enhancing the descriptive power of cognitive assessments of reading, mathematics and science. The item parameters, inter-coder reliability and scaling outputs of PISA-D are comparable to main PISA studies. The risk that some of the constructs used in PISA cannot be adequately applied, adapted or operationalised in some LMIC contexts to ensure international comparability proved to be manageable.

4. PISA contextual questionnaires: questionnaires for students, teachers, schools, parents and system-level data collection enhanced

The focus of the PISA contextual questionnaires is on understanding how measures of student performance at age 15 are related to various aspects of school and classroom practice as well as other related factors, such as economic, social and cultural context. The PISA-D questionnaires include these aspects and also cover a broader set of well-being outcomes and a wider range of risk and protective factors, taking into account differences in life experiences of children in LMIC, both of those who are in school and of those who are not.

A review of the experience of LMIC participating in PISA 2000 to 2015 shows that the PISA questionnaires do not always capture the most relevant contextual factors for these countries. For example, questions about school infrastructure and teaching and learning materials are related to student performance in high-income countries, but are often unrelated to differences in performance in middle-income countries (Lockheed, Prokic-Bruer and Shadrova, 2015[1]). In addition, the measure of economic, social and cultural status used by PISA does not adequately capture lower levels of parental education and income or the risk factors associated with poverty that are more frequent in LMIC.

PISA-D enhances the contextual questionnaires to better measure factors that are more strongly related to student performance in LMIC, while maintaining comparability with PISA on a set of core indicators. For example, the questionnaires collect more detailed data on students' language of instruction at school, language at home and their family socio-economic status, as measured by home possessions and parents' education, literacy skills and participation in the labour force. The questionnaires also identify additional indicators of educational success beyond performance on the PISA test. These indicators comprise, for example, questions about educational attainment, health and well-being, and attitudes towards school and learning.

In addition to assessing student performance, PISA-D introduces an out-of-school assessment to collect data on youth who have not been eligible to sit the PISA school-based test. The out-of-school instruments gather much of the same data as the school-based instruments, as well as data on barriers to school attendance and factors that may impede students' progress through school. It is important to note that in the case of the out-of-school assessment, the data are collected on a tablet computer and through an interview carried out in the household as opposed to a paper-based questionnaire completed by the student in a school setting.

4.1. The contextual framework for PISA-D

The PISA-D questionnaire framework uses the Education Prosperity model (Willms, 2015[3]) as an overarching framework, while also taking into account the goals of PISA-D, lessons from past PISA cycles and other international studies, recommendations from research literature and the priorities of the participating countries. Education prosperity, as applied in PISA-D, is a life-course approach that includes a core set of metrics for success at six key stages of development, covering the period from conception to adolescence. It identifies a key set of outcomes called "Prosperity Outcomes" for six stages of development from conception to age 18, and a set of family, institutional and community

factors, called "Foundations for Success", which drive these outcomes. PISA-D focuses on the fifth stage of the Educational Prosperity framework, late primary and lower secondary (ages 10 to 15).

The framework places great emphasis on equality and equity, with, as noted in an earlier chapter, equality referring to differences among sub-populations in the distribution of their educational outcomes and equity referring to differences among sub-populations in their access to the resources and schooling processes that affect schooling outcomes. The PISA-D contextual framework also focuses on the measurement of socio-economic status and poverty, with the purpose of exploring an international measure of poverty for youth in LMIC; while also extending the measure of the PISA index of economic, social and cultural status (ESCS).

The framework for the PISA-D questionnaires focuses on 15 modules of content. These modules measure the four Prosperity Outcomes, the five Foundations for Success, and the six demographic factors relevant to assessing equality and equity that are listed in Table 4.1. In addition, the questionnaires include several teacher, school and system-level background measures that provide context for the Prosperity Outcomes.

Table 4.1. Modules assessed in the PISA-D questionnaires

1. Prosperity Outcomes	1.1 Academic performance (measured through the PISA-D tests)
	1.2 Educational attainment
	1.3 Health and well-being
	1.4 Attitudes towards school and learning
2. Foundation of Success	2.1 Inclusive environments
	2.2 Quality instruction
	2.3 Learning time
	2.4 Material resources
	2.5 Family and community support
3. Demographic factors for assessing equality and equity	3.1 Gender
	3.2 Socio-economic status and poverty
	3.3 Language spoken at home and language of instruction
	3.4 Urban/rural status
	3.5 Immigrant status
	3.6 Disability

PISA-D enhances the contextual questionnaires to better measure factors that are more strongly related to student performance in LMIC, while maintaining comparability with PISA on a set of core indicators. For example, the questionnaires collect more detailed data on students' language of instruction at school, language at home and their socio-economic status, as measured by home possessions and parents' education, literacy skills and participation in the labour force. The questionnaires also identify additional indicators of educational success beyond performance on the PISA test. These indicators are measured through questions about educational attainment, health and well-being, and attitudes towards school and learning.

The contextual information collected through the questionnaires comprises only a part of the information available to PISA-D. System-level data describing the general structure of the education systems was also used in the PISA-D analysis and country reports. This system-level data includes information on the structure of national programmes, national assessments and examinations, instruction time, teacher training and salaries, educational finance (including enrolment), national accounts and population data.

Available data on all of these indicators have been reviewed for PISA-D countries, identifying the current status of system-level data collection and availability in terms of quality and completeness (UNESCO Institute for Statistics, 2016[4]).

4.1.1. The PISA-D school-based questionnaires

The school-based questionnaires for students, teachers and the principals of schools have been developed in accordance with the contextual framework. These questionnaires take about 35 minutes for the students to complete and about 25 minutes each for teachers and the principals. The responses to the questionnaires are analysed with the assessment results to provide at once a broader and more nuanced picture of student, school and system performance. These questionnaires seek information about:

- students and their family backgrounds, including their economic, social and cultural capital, and the language they speak at home versus the language of instruction

- aspects of students' lives, such as their level of educational attainment, their health and well-being, and their engagement with school

- aspects of learning, including quality of instruction, inclusive environments, learning time, school material resources and family and community support

- contexts of learning, including teacher, school and system-level information.

4.1.2. The PISA-D out-of-school questionnaires

The out-of-school component questionnaires for youth, parents and interviewers have been developed in accordance with the contextual framework. These questionnaires take between 15 and 30 minutes each for the youth, the person most knowledgeable about the youth (parent, guardian or other) and the interviewer to complete. These questionnaires seek information about:

- youths and their family backgrounds, including their economic, social and cultural capital, and the language they speak at home versus the language of instruction when they attended school

- aspects of youths' lives, such as their level of educational attainment, their attitudes towards learning, their employment status, their habits and life outside of school, and their health and well-being

- aspects of learning, including inclusive environments, family support, their perception of the inclusiveness of their school environment when they attended school, their reasons for being out of school and barriers preventing them from returning to school, and their family support and environment

- aspects of youths' early years, their educational experience and their parent or care-giver's educational expectations for the youth

- aspects of youths' households, including location and surrounding characteristics.

The distribution of questions across the elements of the Educational Prosperity model is shown in Figure 4.1, with red circles indicating questions that can be linked with PISA 2015 and blue circles indicating questions that are new to PISA-D.

Figure 4.1. Questions in the PISA-D contextual questionnaires

	School-based assessment			Out-of-school assessment		
	Student	Teacher	School	Youth	Person most knowledgeable about the youth	Household
Prosperity Outcomes						
Educational attainment	🔴🔴🔴🔴			🔴🔴🔵🔵🔵🔵🔵🔵🔵 🔵🔵🔵🔵🔵	🔵🔵🔵	
Health and well-being	🔴🔵🔵🔵			🔴🔵🔵🔵		
Attitudes towards school and learning	🔴			🔵🔵	🔵	
Foundations for Success						
Inclusive environments	🔴🔵🔵🔵	🔵🔵	🔴🔵🔵🔵	🔵🔵🔵		
Quality instruction	🔴🔵🔵	🔵	🔵			
Learning time	🔴🔵🔵🔵🔵	🔵	🔴🔵			
Material resources		🔵🔵🔵🔵🔵🔵	🔵🔵🔵🔵🔵			
Family and community support	🔵	🔵	🔵🔵	🔵	🔵	
Demographic factors to assess equity and equality						
Gender	🔴			🔴		
Socio-economic status and poverty	🔴🔴🔴🔴🔴🔴🔴🔴 🔵🔵🔵🔵🔵🔵🔵🔵🔵 🔵🔵		🔵🔵	🔴🔴🔴🔴🔴🔴🔴🔴🔴 🔴🔵🔵🔵🔵🔵🔵🔵🔵 🔵🔵🔵🔵🔵🔵🔵🔵🔵 🔵🔵🔵🔵🔵🔵🔵🔵🔵 🔵🔵	🔵🔵🔵🔵🔵🔵🔵🔵	🔵🔵🔵🔵🔵🔵🔵
Language spoken at home	🔴🔵🔵	🔵🔵🔵		🔴🔵🔵🔵		
Urban/rural status			🔴			
Immigrant status	🔴			🔴🔴🔴		
Disability	🔵			🔵🔵		
Context factors		🔴🔴🔴🔴🔴🔴🔵🔵🔵🔵 🔵🔵🔵🔵🔵🔵🔵🔵	🔴🔴🔴🔴🔴🔴🔴 🔵🔵🔵🔵		🔴🔴🔴🔵🔵🔵	🔵🔵🔵🔵🔵🔵🔵
Total	49	33	28	77	19	14

Note: Red dots indicate questions that come from the main PISA assessment; blue dots indicate questions new to PISA-D.

4.2. Student and out-of-school youth contextual questionnaires

The PISA contextual questionnaires focus on understanding how 15-year-old students' performance is related to certain school and classroom practices, and to other factors, such as students' economic, social and cultural background. The PISA-D student questionnaire includes many of the questions found in the student questionnaire in the main

PISA assessment and also covers a broader set of well-being outcomes and a wider range of risk and protective factors related to performance. In addition to the student questionnaire, PISA-D introduces a questionnaire for 14-16 year-olds who are not enrolled in school or in PISA's target grades (grade 7 or above). The out-of-school youth questionnaire gathers much of the same data as the student questionnaire and adds questions about why the child is not in school, barriers preventing the child from returning to school, and employment.

Information collected through the PISA-D student and out-of-school youth questionnaires includes, but is not limited to, the following:

- *Educational attainment*: Both PISA and PISA-D collect data on student's grade, whether they have repeated a grade and whether they have attended early childhood education. PISA-D also asks out-of-school youth about whether they work, their profession, hours worked per week, and their wage or salary.

- *Health and well-being*: Both PISA and PISA-D ask about general life satisfaction. Taking a deeper look at health and well-being, PISA-D asks about young people's general perception of their health; their physical and mental health during the past year; and emotional distress, including anxiety and depression.

- *Student engagement*: Like PISA, PISA-D includes a measure of general attitudes towards school and learning. Out-of-school youth are expected to respond to these questions based on their experiences when they last attended school.

- *Inclusive environments*: Similar to PISA, PISA-D asks students to report on their sense of belonging at school. It further explores school climate with questions on whether students feel safe at school and whether they have been sexually harassed at school. Out-of-school youth respond based on their experiences when they last attended school.

- *Quality instruction*: Both PISA and PISA-D assess the learning climate in the classroom and teacher-student relations. PISA-D adds new questions on the structure of the classroom and teaching practices in the student's mathematics lessons.

- *Learning time*: Like PISA, PISA-D asks students about loss of learning time due to truancy. PISA-D explores learning time in greater depth by asking both students and out-of-school youth about long-term absenteeism and reasons for missing school for long periods. It further asks about reasons for reduced teaching time, the time students take to travel from their home to school, and reasons why out-of-school youth do not participate in schooling.

- *Family and community support*: PISA-D asks about the types of communication students and out-of-school youth have with their parents or other family members.

- *Demographic factors*: PISA and PISA-D gather information on students' *gender* and *immigrant background* (based on answers to questions of where students and their parents were born). PISA-D is the first PISA study to include self-reported measures pertaining to *disability*. PISA-D also extends PISA's measure of *socio-economic status* to capture risk factors of *poverty* by expanding PISA's question on home possessions and gathering more details on parents' education and profession. While PISA asks students about the *language spoken at home*, PISA-D

asks about languages spoken at home and by instructors at school, and the language they first learnt to read.

4.2.1 Examples of new questions designed for the PISA-D student and out-of-school youth contextual questionnaires

The following questions are asked to both students and out-of-school youth to discern information about long-term absenteeism and the reasons for it.

ST011 ST011Q01NA	**Have you ever missed school for more than three months in a row?** *(Please tick only one box.)*		
	No	$\square_1$	**Please go to Question 18.**
	Yes, once	$\square_2$	**Please go to Question 17.**
	Yes, twice or more	$\square_3$	**Please go to Question 17.**

ST012	**Why did you miss school for more than three months in a row?** *(Please tick one box in each row.)*	Yes	No
ST012Q01NA	I was bored.	$\square_1$	$\square_2$
ST012Q02NA	I was suspended for something I did (e.g., violence, aggression, use of drugs, drug dealing).	$\square_1$	$\square_2$
ST012Q03NA	I was pregnant.	$\square_1$	$\square_2$
ST012Q04NA	I could not reach school because of transportation problems.	$\square_1$	$\square_2$
ST012Q05NA	I did not have a teacher.	$\square_1$	$\square_2$
ST012Q06NA	I could not understand the language in which the lessons were given.	$\square_1$	$\square_2$
ST012Q07NA	I had to take care of a family member.	$\square_1$	$\square_2$
ST012Q08NA	I had to help with work at home or on the family land.	$\square_1$	$\square_2$
ST012Q09NA	I had to get work to bring money home.	$\square_1$	$\square_2$
ST012Q10NA	I was sick.	$\square_1$	$\square_2$
ST012Q11NA	I did not feel safe at school.	$\square_1$	$\square_2$
ST012Q12NA	I had to take care of sick parents or relatives.	$\square_1$	$\square_2$
ST012Q13NA	I was no longer interested in school.	$\square_1$	$\square_2$
ST012Q14NA	I could not pay <school fees>.	$\square_1$	$\square_2$
ST012Q15NA	School was closed because of a natural disaster (e.g., flood, earthquake).	$\square_1$	$\square_2$

The following questions are asked to both students and out-of-school youth to gather information on lower levels of socio-economic status. These are an example of how PISA-D extends the PISA index of economic, social and cultural status (ESCS) to capture lower levels of income and wealth in LMIC.

ST051 ST051Q01NA	**What is the floor of your home mostly made of?** *(Please tick only one box.)*	
	Earth, sand or dung	$\square_1$
	Wood planks, palm, or bamboo	$\square_2$
	Parquet, polished wood	$\square_3$
	Vinyl or asphalt strips	$\square_4$
	Ceramic tiles	$\square_5$
	Cement	$\square_6$
	Stone	$\square_7$

ST059 ST059Q01NA	**In the past 30 days, how often were you hungry because there was not enough food?** *(Please tick only one box.)*			
	Never or almost never	*About once a week*	*2 to 3 times a week*	*Almost every day*
	$\square_1$	$\square_2$	$\square_3$	$\square_4$

The following question is asked to out-of-school youth to help understand reasons why they do not participate in schooling.

YI019c	**We would like to know whether certain factors would enable you to continue with your schooling. Please listen as I read a list of factors and tell me whether any of these factors would enable you to continue your schooling.** **Would you be more likely to continue your schooling if there were …..** *[INTERVIEWER: Read each item separately and indicate Yes or No (or Don't know or Refuse). If required, repeat the stem, "Would you be more likely to continue your schooling if there were …"]*				
		Yes	**No**	**DK**	**RF**
Y019CQ01NA	A school that was closer to home	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ02NA	Better quality teaching	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ03NA	Accessible for students with disabilities	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ04NA	A safe school	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ05NA	No discrimination, such as gender, racial, ethnic, religious discrimination	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ06NA	A school that was more accepting of students' differences	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ07NA	Financial incentives, such as conditional cash transfers	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ08NA	No school fees or if I did not have to pay them	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ09NA	A vocational program that would result in a job	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ10NA	Help with improving my reading skills	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ11NA	A school where I could learn at my own pace	$\square_1$	$\square_2$	$\square_3$	$\square_4$
Y019CQ12NA	Support for my children while attending school	$\square_1$	$\square_2$	$\square_3$	$\square_4$

4.3. Teacher and school contextual questionnaires

In addition to the student questionnaire, the school-based component of PISA-D includes questionnaires for teachers and the school administration. The PISA-D teacher and school questionnaires include many of the questions from the main PISA assessment but enhance these instruments by covering a wider range of risk and protective factors related to student performance. PISA-D's teacher questionnaire asks teachers to provide information about themselves, their school's resources, their teaching practices and their students. PISA-D's

school questionnaire asks school principals about their school, its students and teachers, and the learning environment.

The PISA-D school and teacher questionnaires collect data on the Foundations for Success, and provide some detail about the demographic factors for assessing equality and equity. The questionnaires also include several teacher, school and system-level background measures, many of which come from PISA 2015, providing additional context for the Prosperity Outcomes. Information collected through the teacher and school questionnaires includes, but is not limited to, the following:

- *Inclusive environments*: As with PISA, PISA-D asks principals about school policies concerning how students are admitted to the school and grouped for instruction, and about the diversity of the school. PISA-D also asks school principals about their attitude towards grade retention and asks teachers about their attitude and practices towards teaching less able students.

- *Quality instruction*: PISA-D asks teachers about their practices for teaching less able students. PISA-D also asks school principals about teachers' behaviours that could negatively impact classroom climate and the quality of instruction delivered to students at their school.

- *Learning time*: PISA-D asks school principals about the reasons for and the amount of instructional time lost during the last year, as does PISA. PISA-D adds questions to teachers about the reasons they are absent and to principals about their policies regarding teacher absenteeism.

- *Material resources*: PISA-D asks teachers about the availability and condition of school resources as well as the teacher's use of those resources. PISA-D questions to school principals focus on the availability and condition of school infrastructure and facilities as well as the availability of textbooks, whereas PISA asks whether the school lacks resources and collects information on the availability of ICT resources and internet connectivity.

- *Family and community support*: PISA-D asks teachers about families' involvement at school and asks school principals about how parent and community members or organisations contribute to the school, whereas PISA gathers information about school policies for parental involvement.

- PISA-D's six demographic factors for assessing equality and equity are reported on primarily through the student and out-of-school youth questionnaires. However, the PISA-D school questionnaire, like PISA, collects information about school meals as part of the measure of *socio-economic status*; and PISA-D asks teachers about the *language of instruction* versus the language they use when talking with students.

PISA-D also includes school and system-level background measures that can help explain student outcomes. Some are from PISA, e.g. teacher professional development activities and school resources, management, grade retention policies and academic support services. PISA-D adds new questions about whether the teacher teaches multi-grade classes, holds multiple teaching jobs or works other jobs in addition to teaching, school location and nearby hazards. Also unique to PISA-D, teachers are asked about the proportion of students in their class that lack the necessary literacy skills to acquire the knowledge intended in the curriculum.

4.3.1 Examples of new questions designed for the PISA-D teacher and school contextual questionnaires

The following question about inclusive environments is asked to teachers to gather information about their attitudes towards dealing with less able students.

TC028 — **To what extent do you agree with the following statements?**
(Please tick one box in each row.)

		Strongly Disagree	Disagree	Agree	Strongly Agree
TC028Q01NA	Many of the students who are behind should have been <held back>.	☐1	☐2	☐3	☐4
TC028Q02NA	Students who repeat one or more <grades> will eventually leave school before they graduate.	☐1	☐2	☐3	☐4
TC028Q03NA	After students have repeated a <grade>, they are still unprepared for the next grade.	☐1	☐2	☐3	☐4
TC028Q04NA	Students who have repeated a <grade> learn the hard lesson that they must try harder to succeed.	☐1	☐2	☐3	☐4
TC028Q05NA	Teachers should try to teach the curriculum, even to students who do not have the basic reading and numeracy skills.	☐1	☐2	☐3	☐4
TC028Q06NA	Students with disabilities should be taught in <special schools>.	☐1	☐2	☐3	☐4
TC028Q07NA	Teachers waste their time trying to support teen mothers to remain in school.	☐1	☐2	☐3	☐4
TC028Q08NA	Teachers should adjust the curriculum to the cultural diversity in their classes	☐1	☐2	☐3	☐4
TC028Q09NA	Students who lag behind should be placed in special classes.	☐1	☐2	☐3	☐4
TC028Q10NA	Teachers should be able to teach classes with students with differing levels of ability.	☐1	☐2	☐3	☐4
TC028Q11NA	A student should never have to repeat more than one <grade>.	☐1	☐2	☐3	☐4

The following question about material resources is asked to principals to determine the availability and condition of school facilities.

SC012 — **Does your school have the following features, and if so, in what condition are they?**
(Please tick one box in each row.)

		No, not available	Yes, but in poor condition	Yes, but in need of minor repairs	Yes, in good condition
SC012Q01NA	Flush toilets	☐1	☐2	☐3	☐4
SC012Q02NA	Other types of toilets (e.g., <latrines>, <squat holes>, <pit toilets>)	☐1	☐2	☐3	☐4
SC012Q03NA	Kitchen	☐1	☐2	☐3	☐4
SC012Q04NA	Place with drinkable water	☐1	☐2	☐3	☐4
SC012Q05NA	Running water	☐1	☐2	☐3	☐4
SC012Q06NA	Electricity	☐1	☐2	☐3	☐4
SC012Q07NA	Indoor plumbing	☐1	☐2	☐3	☐4
SC012Q08NA	<First aid room>	☐1	☐2	☐3	☐4
SC012Q09NA	Immunisation or health care room	☐1	☐2	☐3	☐4
SC012Q10NA	Cafeteria	☐1	☐2	☐3	☐4
SC012Q11NA	Sports area or playground	☐1	☐2	☐3	☐4
SC012Q12NA	Fence or hedge on the school borders	☐1	☐2	☐3	☐4
SC012Q13NA	<Access ramp>	☐1	☐2	☐3	☐4
SC012Q14NA	<Fans>	☐1	☐2	☐3	☐4
SC012Q15NA	<Lighting>	☐1	☐2	☐3	☐4

4.4. Out-of-school youth's parent contextual questionnaire

For its out-of-school assessment, PISA-D introduces out-of-school youth, parent and household observation questionnaires. The PISA-D questionnaire administered to the parent or care-giver of the out-of-school youth collects information on the youth's life and their family background.

The PISA-D out-of-school youth's parent questionnaire collects data on three of the Prosperity Outcomes, one of the Foundations for Success, and provides detail about the demographic factors, socioeconomic status and poverty. It also includes some background measures, which provide additional information about the person responding to the questionnaire. This questionnaire offers a unique opportunity to ask about the youth's prenatal and early life experiences. Information collected through the parent questionnaire includes the following:

- *Educational attainment*: PISA-D asks the parent or care-giver whether the youth received formal early childhood and care and whether the youth attended pre-primary education, and also asks about their educational expectations for the youth as well as factors that could hinder the youth's completion of compulsory education.

- *Student engagement*: PISA-D asks the parent or care-giver about their values towards schooling outcomes and attitudes towards education.

- *Health and well-being*: PISA-D asks the parent or care-giver about the youth's prenatal and early life experiences related to health, such as the mother's health during pregnancy, the conditions and any complications of the mother's birth experience, how the youth was fed during the first six months of life, whether the youth had any health problems during the first five years of life, and whether the youth received vaccinations.

- *Family and community support*: PISA-D asks the parent or care-giver about the type of support they provided to the youth in their early years.

- PISA-D's six demographic factors for assessing equality and equity are reported on primarily through the student and out-of-school youth questionnaires. However, the PISA-D parent questionnaire collects information about the out-of-school youth's early childhood living conditions as part of the measure of *socioeconomic status and poverty* with questions about food security during their first two years of life and whether they received government support for schooling.

PISA-D also includes some background measures about the person responding to the parent questionnaire, including their relationship to the youth, occupation and highest level of education, and whether they had help completing the questionnaire.

As the PISA-D out-of-school component is conducted as a household survey, an interviewer is required to administer both the out-of-school youth assessment and the parent questionnaire. The interviewer completes a brief household observations questionnaire at the end of the interview, providing further background information on the youth interview process, including who was present during the interview, whether they assisted the respondent in answering the parent questionnaire, whether they assisted the youth in answering the cognitive test, whether the youth understood the interview questions and where the interview took place. The household observations questionnaire also collects data contributing to the measure of *socioeconomic status and poverty* with questions about the

youth's type of dwelling and location, discerning whether it is in a rural or urban setting, and about the area surrounding the household, i.e. the type of road leading to the dwelling, the presence or absence of street lights surrounding the dwelling, security concerns, roof and wall material of the dwelling, and the presence or absence of electricity.

4.4.1 Examples of new questions designed for the PISA-D out-of-school youth's parent contextual questionnaire

The following question asked to parents of out-of-school youth provides insight to the extent to which the parent values schooling outcomes.

HH020	**Thinking about school, to what extent do you agree with the following statements?** *(Please tick one box in each row.)*	Strongly agree	Agree	Disagree	Strongly disagree	I don't know	I refuse to answer
HH020Q01NA	School does little to prepare youth for adult life	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q02NA	School helps getting a job	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q03NA	School is a waste of time	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q04NA	School teaches how to be a citizen	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q05NA	School ignores native languages and cultures	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q06NA	School is a waste of money	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q07NA	School integrates people into society	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q08NA	School builds confidence to make decisions	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q09NA	School teaches useful work skills	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q10NA	Top-of-their class graduates get very good jobs	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q11NA	School is the only way to get better life opportunities	☐1	☐2	☐3	☐4	☐93	☐94
HH020Q12NA	School helps overcoming ignorance	☐1	☐2	☐3	☐4	☐93	☐94

4.5. System-level data

In addition to contextual data collected from students and schools, the collection of system-level data are an important part of PISA that facilitates the analysis and interpretation of test results. Like PISA, PISA-D collects system-level data. Each country's PISA-D National Centre also completes a system-level data questionnaire on the structure of the country's education system, assessments and examinations, instruction time, teacher training and salaries, national accounts, expenditure on education and enrolment. All data collected helps analysts interpret the assessment results. The system-level data also helps OECD and non-OECD Members (Partners) understand how elements of an education system interact with each other to either secure or impede better learning outcomes. The data also describe the contexts in which education systems function and the policy dynamics operating within those systems. With this information, countries can monitor progress towards improvement targets, analyse factors associated with outcomes, and build institutional capacity for managing large-scale assessments and evidence-based policy making.

The UNESCO Institute of Statistics (UIS) was commissioned jointly by the OECD and the World Bank, a PISA-D development partner, to review the status of system-level data collection and availability (quality and completeness) in six PISA-D countries (UNESCO

Institute for Statistics, 2016[4]). UIS identified country-specific issues to be addressed as part of project implementation, as system-level data are used in analysing and reporting PISA-D results. UIS also provided lessons learnt for working with increasing numbers of LMIC participating in future PISA cycles.

Six of the nine PISA-D countries (Cambodia, Ecuador, Guatemala, Paraguay, Senegal and Zambia) completed the system-level data questionnaire in 2015 during the design, planning and co-ordination phase of the project. Honduras joined the project at the end of 2015 and is completing the system-level data questionnaire in 2017; Panama is participating in PISA2018 in addition to the out-of-school component of PISA-D, so its system-level data are collected through PISA; and Bhutan joined in 2017 to participate in the main data collection using the cognitive-assessment instruments only. Overall, system-level data collection has shown that PISA-D countries can provide the requested data, and that quality data and metadata are available or can be produced with some additional work. It has also shown that the six countries:

- can provide information on the **structure of their education system**, the theoretical age at entry, and the duration of each education level

- have solid institutions in charge of **national assessments and examinations** at lower and upper secondary level

- have national frameworks and implementation guidelines for **instruction time** in public institutions, by grade

- regulate **pre-service teacher training**, **entry into the profession** and **professional development**, but in several countries, teacher training standards and policies are changing

- can produce data on **expenditure** on and **enrolment** in public educational institutions.

Although each country faces its own challenges in collecting and reporting on education system-level data, certain issues have emerged as common to a majority of the countries:

- Countries lack accurate data on private expenditure on education.

- There can be discrepancies between policy and actual practice, as countries sometimes lack control or disciplinary mechanisms to ensure that practice adheres to national standards.

- Education system data are not available in a single hub, but has to be collected from numerous databases maintained by different ministries and agencies.

Some of these challenges might also be faced by countries now participating in PISA. The UIS review of PISA-D system-level data collection produced recommendations for future PISA cycles:

- Countries have made concerted efforts to improve national standards, such as by upgrading minimum teaching standards. As a result, assessing the teaching force according to academic qualifications requires close scrutiny of official records.

- As different institutions within each country are responsible for the various elements covered by the system-level data questionnaire, the PISA-D national centre must communicate and co-ordinate with each of these institutions.

- The OECD could consider modifying the system-level data questionnaire to avoid collecting data and metadata that are already reported and available in international databases.

The OECD has taken these recommendations on board in its preparation for PISA 2018, particularly with regard to the system-level data questionnaire.

4.6. Findings from the PISA-D field trial of the school-based contextual questionnaires

The PISA-D school-based contextual questionnaires were piloted in seven countries: Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia. The field trial of the school-based assessment instruments took place from August to December 2016. Based on results of the field trial, the instruments and survey operations were modified as necessary for the Main Survey.

The PISA-D school-based assessment field trial revealed that the instruments work as intended. Findings from the analysis of the PISA-D contextual questionnaire field trial data are organised around four major goals and are summarised in the table below.

Table 4.2. Findings from the field trial of the PISA-D cognitive test

Contextual questionnaire field-trial goals	Findings
Select items that provide reliable measures of the core indicators (achievement, attainment, health/well-being, student engagement, inclusive environments, quality instruction, learning time, material resources, family/ community support, gender, disability, immigrant background, socio-economic status/poverty, language spoken at home/school)	The first priority for selecting content was to identify the questions relevant to the core indicators. The analysis showed reliable measures for these. Analyses at the question level were conducted in tandem with analyses of items within scales. Items were chosen based on their psychometric properties, including their role in the factor structure, contribution to reliability of the scale, their item-response theory parameters (difficulty and discrimination), and their consistency across cultures (e.g. differential item functioning).
Include measures that can be used to link to PISA 2015	The second priority was to select linking measures that were included in PISA 2015. Around 40% of the PISA-D student questions link to PISA, as do 20% of the teacher questions and 40% of the school questions.
Provide the constituent components to measure socio-economic status and poverty	Analyses of data from seven countries showed that the selected items provide a reliable measure for the extended measure of home possessions and a new measure of poverty.
Include a number of measures of supporting content (teacher, school and system-level background variables that help explain student outcomes)	A number of questions were included in the questionnaires that were intended to provide supporting content for the core indicators. The results of the field trial facilitated decisions on which of these supporting questions should be retained or deleted to satisfy space limitations in the Main Survey.

As noted in the previous chapter, the results of the field trial also confirmed that the participating countries carried out their survey operations in accordance with PISA's technical standards, and that their data are internationally comparable. All findings from the PISA-D field trial, including lessons learnt, were documented for reference in preparing for the main data collection.

4.7. Findings from the PISA-D Main Survey of the school-based contextual questionnaires

As mentioned in the previous chapter, the main survey data collection is subject to a strict adjudication process, and the PISA-D adjudication review suggests that the PISA-D contextual questionnaires have been implemented in accordance with PISA's technical

standards and that the isolated deviations from these standards (listed in the previous chapter) did not affect the quality of the data products. The quality with which procedures were implemented in PISA-D is in fact comparable to any previous cycle of PISA.

During the analysis phase of the project, analysts took an in-depth look at the results and confirmed that the instruments measure what they purport to measure against the *PISA-D Assessment and Analytical Framework* (OECD, 2018[5]).

4.7.3 International comparisons

PISA-D was successful in making the assessment instruments more relevant to LMIC while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA. These variables and factors include student performance, educational attainment, health and well-being, attitudes towards school and learning, the learning environment, learning time (e.g. Figure 4.2), the quality of instruction, family and community support, and resources devoted to education.

Figure 4.2. Loss of learning time, PISA 2015 and PISA-D

Percentage of students who reported that they skipped school, skipped some classes or arrived late for school at least once in the two weeks prior to the PISA test; and percentage of students reporting if they have ever missed school for more than 3 months in a row (unique to PISA-D)
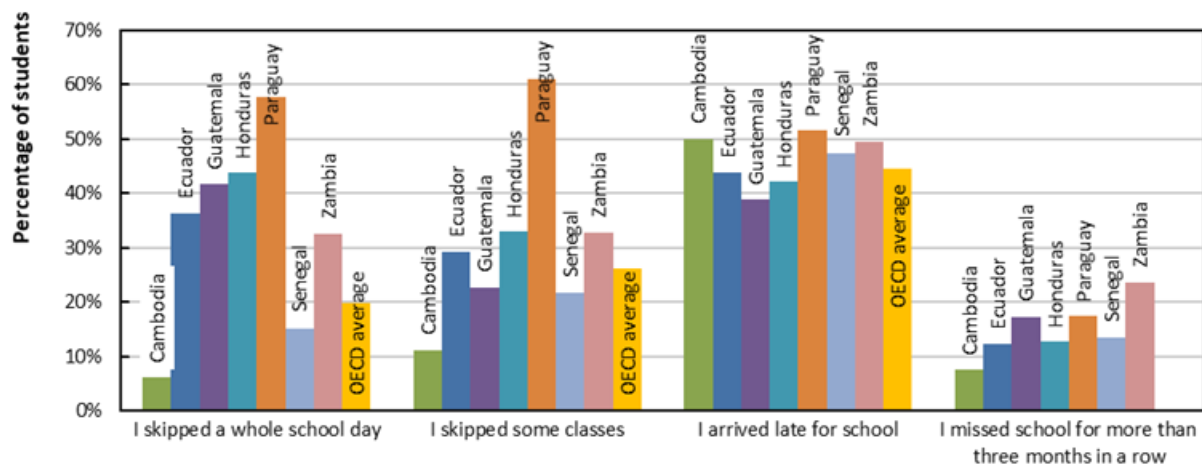
Source: PISA 2015 Database and PISA for Development Database.

4.7.4 Policy insights

Students' performance on the PISA-D tests at age 15 is the result of an accumulation of various factors that affect children's development, beginning at conception and continuing through to the time of the assessment. Therefore, caution is advised when considering, for example, whether school or classroom practices, gleaned from responses to PISA-D contextual questionnaires, have strong relationships with performance. However, it is possible to identify a range of factors that influence student performance and related outcomes, and PISA-D refers to these as Foundations for Success. PISA-D results reveal a need across all countries to improve:

- the allocation of resources in education (i.e. eliminate grade repetition, allocate resources equitably, improve the school environment, and reduce student truancy and teacher absenteeism)

- the school environment (i.e. provide an environment conducive to learning, ensure educators set the tone for an inclusive school community)

- the quality of instruction (i.e. make teaching more effective)

- family and community support for education.

The countries participating in PISA-D benefit by understanding how well their students fare compared with students in other countries. The results of PISA-D allow participating countries to determine whether their policies differ from those of countries with a similar social and economic context, but whose students perform better and benefit from more equitable learning opportunities. These comparisons can provide valuable peer learning, and can help strengthen a country's political will to invest resources in education and identify effective policies that they can adapt to their particular contexts.

4.8 Achievement of the output

This output of the project has been achieved on time and within the scope of the project budget. Questionnaires and data collection instruments were developed and trialled on schedule. The item parameters, inter-coder reliability and scaling outputs of PISA-D are comparable to main PISA studies. The risk that some of the constructs used in PISA cannot be adequately applied, adapted or operationalised in some LMIC contexts to ensure international comparability proved to be manageable.

5. Methodologies and approaches for incorporating out-of-school youth into the assessment

In most OECD countries and across many other PISA-participating countries and economies, enrolment in school at age 15 is nearly universal, and schooling is compulsory until approximately that age. However, in many LMIC – including some of those that have participated in PISA – relatively large proportions of 15-year-olds are not enrolled in school or are not enrolled in PISA's target grades (grade 7 and above). With an increasing number of LMIC participating in PISA, and with 61 million children of lower secondary school age out of school around the world, the OECD set out with PISA-D to ensure that this population is no longer beyond the reach of programmes that try to evaluate the success of education systems.

The denial of education continues to be shaped by wealth, with significant gaps between out-of-school rates in the world's richest and poorest countries. These gaps begin to appear at the primary level, with almost every child in primary school in high-income countries, while this is just the case for 80% in LMIC. And the gaps widen with age: 60% of youth of upper secondary school age are not in school in LMIC, compared to just 6% in high-income countries.

A key rationale for the PISA-D out-of-school assessment pilot is that measures of school achievement through administration of tests refer to students - not the whole population – and this poses a problem for assessing efficiency and human capital, especially in LMIC. With regard to the efficiency of an educational system, it is important to see test scores in the context of participation rates and the extent to which the dropout rate in a country has been reduced between any two measurement points. With regard to evaluating human capital – the whole population – it is essential to know the human capital of school dropouts and those that never enrolled. In addition, dropout and never enrolling is very negatively correlated with socio-economic status. An assessment that ignored the out-of-school is at risk of perversely encouraging policies of exclusion.

While it is possible to estimate test scores for the whole population (i.e. taking into account dropouts) by putting bounds on unobserved scores, this is effectively guess-work and is carried out under weak assumptions. There is no substitute for assessing the skills of the whole population.

5.1. PISA for Development piloted an out-of-school assessment for use in future PISA cycles

As part of its efforts to make PISA more relevant to LMIC, the OECD piloted an approach to include out-of-school youth in the assessment. PISA-D collected data on the skills and non-cognitive attributes of 14-16-year-olds who are not included in PISA's in-school assessment. The aim was to gather better, actionable data on the characteristics of this population, the reasons why these children are not in school and on the magnitudes and types of exclusion and disparities in access to education and learning opportunities.

PISA-D countries Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia piloted the out-of-school assessment, but only the first five of these countries proceeded to the Main Survey data collection phase. In order to reach its targeted population, the assessment included background questionnaires and test questions that are relevant to out-of-school

youth and mechanisms to sample, identify and contact them. As noted above, the PISA-D out-of-school assessment instruments were delivered on tablet computers during household surveys conducted by trained interviewers. As already noted earlier in this report, the respondent completed the test on the tablet computer after a brief introduction by the interviewer. The respondent completed the questionnaire with the support of the interviewer and the interviewer completed a household observation questionnaire.

The range of educational experiences in the out-of-school population was expected to vary substantially, from children with no experience in formal education to those who have recently left school or who are still in school but in grade 6 or below. Therefore, the assessment incorporates a set of screening questions that allows the interviewers to provide children with the test forms that are most appropriate to their level of education. The background questions also aim to elicit responses that will help analysts identify the barriers to attending school, ways to prevent dropout, and the obstacles to children's smooth progression through formal schooling.

5.1.1. PISA-D out-of-school results are comparable to PISA-D in-school results and to the main PISA results

The in-school and out-of-school PISA-D assessments share the same framework, focusing on the lower end of the performance spectrum and emphasising lifelong learning, access to schooling, and equity in education opportunities. Many of the cognitive test items and background questionnaire items overlap with those used in the paper-based PISA-D in-school assessment. However, out-of-school youth are assessed in reading and mathematics only.

The scientific literacy domain was not included in the out-of-school assessment due to practical considerations of total assessment time and the burden on individuals in a household survey. On one hand, the total test allowed a maximum of 50 minutes, which did not provide adequate time to include an assessment of three domains and meet the requirements of linking to the PISA scales. Therefore it became necessary to choose only two domains. In making the decision, it was taken into account that reading and mathematics literacy are considered foundational skills that are necessary for the development of scientific literacy skills. In addition, the target population was also taken into account. As science is the domain with the strongest link to school-based learning, this domain was the least appropriate for a group that, by definition, has been exposed to less formal schooling. Therefore, it was decided to include reading and mathematics as two domains assessed in Strand C.

The PISA-D out-of-school assessment aimed to yield a large enough sample size to (i) test the validity of the assessment items, including linking results to the PISA-D in-school assessment as well as to the main PISA proficiency scales; and (ii) provide the participating countries with meaningful and policy relevant information about the out-of-school population. The validated instruments, methods and approaches from this pilot project are now being made available as options for all participating countries in future rounds of PISA and as a learning assessment module linked to or integrated with a household survey.

5.2. The PISA-D out-of-school youth assessment

Through the out-of-school assessment attempted in PISA-D, PISA is, for the first time, able to report on what *all* 15-year-olds in a population (in-school and out-of-school) know and can do. The analysis of these data has yielded valuable insights for governments in the participating countries, in particular about the effectiveness of their education systems, and about the success of policies that aim to ensure inclusive and equitable quality education and learning opportunities for all.

5.2.1. The test instrument

As discussed in Chapter 3 of this Report, the PISA-D household-based survey assessed 14-16 year-olds not enrolled in PISA's target grades (grade 7 and above) in the domains of reading and mathematics; each domain was treated equally in the assessment. The PISA-D out-of-school instrument is a tablet-based assessment designed as a 50-minute test. The computer-based household survey Programme for the International Assessment of Adult Competencies (PIAAC) was used as a model for selecting the delivery mode, and tablets were chosen over laptops on account of cost, efficiency and user-friendliness.

The PISA-D out-of-school test includes a ten-minute core module of basic reading and mathematics skills to ensure that respondents have an appropriate level of skills to proceed to the full assessment (see Figure 5.1). An established minimum number of items answered correctly will determine the set of items that will be presented to respondents in the second stage of the cognitive assessment. The second stage was designed to take no longer than 40 minutes to complete. Respondents who pass the core module are randomly assigned to one of the 30 forms measuring reading and mathematical literacy. Respondents who fail the core module are directed to a 10-minute assessment of reading components.

5.2.2. The contextual questionnaires

As discussed in Chapter 4 of this Report, the out-of-school component questionnaires for youth, parents and interviewers were developed in accordance with the contextual framework. The participants in the survey answered a 30-minute questionnaire that was administered via the interviewer on a tablet computer. The questionnaire for the person most knowledgeable about the youth was a paper and pencil based instrument that was left behind by the interviewer in the household and collected later. In addition, the interviewer completed an observation questionnaire on the tablet computer which covered aspects of the youths' households, including location and surrounding characteristics.

5.3. Sampling and survey operations for out-of-school assessment field trial

5.3.1. Sampling

PISA-D's definition of "out-of-school" builds on the work of UNICEF and UIS that has defined those children and young people that are excluded from education opportunities. Based on this definition, youth aged 14 to 16 are included in the PISA-D out-of-school sample if they have never attended school, have attended but dropped out during primary school, have completed primary school but did not continue to secondary school, entered secondary school but dropped out, or are currently enrolled in school but are in Grade 6 or below.

The sample was expanded from 15-year-olds to 14-16 year-olds on the recommendation of OECD Education Working Paper No. 120, referred to in Chapter 2 of this report, which

was drafted by independent expert Roy Carr-Hill and can be found at this link. The working paper highlighted the challenge of locating a single year age group in a household survey in LMIC. Further challenges are presented by the fact that schooling experience in this out-of-school population is likely to vary substantially, from children with no experience in formal education whatsoever, to those who have recently left school or who are still in school but in Grade 6 or below.

Table 5.1. Information about the PISA-D out-of-school target population

		Guatemala	Honduras	Panama	Paraguay	Senegal	Zambia
14-16 years old	Count	1 097 289	569 543	215 101	407 424	953 150	1 066 270
14-16 years old, <7th grade	Percent	13%	5%**	4%	9%	21%	8%
14-16 years old, out-of-school	Percent	33%	43%	13%	17%	32%	21%
	Female	unknown*	49%	48%	46%	43%	54%
	Rural	unknown*	68%	67%	62%	unknown	60%
Overall percentage of 14-16 year-old out-of-school youth in target population		46%	49%	17%***	26%	53%	30-40%***
Areas with relatively higher rates of out-of-school youth		Poorest, rural or no school nearby	Rural areas	Rural areas (because of accessibility problems and non-availability of all school grades)	Occidental region (low density, indigenous population, low coverage of public services) and some departments in the oriental region with a higher rural population	Unknown	Slums in urban areas, trading places in urban areas, rural areas, and streets in urban areas

Note: *Guatemala instead provided the percentage of out-of-school boys (38%) and girls (29%). **Honduras reported 9%, but this is conditional on being enrolled, so it has been adjusted to 5%. ***Percent enrolled >=7th grade, percent enrolled <7th grade, and percent out-of-school do not add up to 100%.
Source: Country responses to an out-of-school population overview questionnaire completed in the beginning of the survey.

The PISA-D out-of-school assessment has three main objectives with regard to sampling:

- To yield a large enough sample to test the validity of the cognitive test, youth questionnaire and other instruments, including linking to PISA and the PISA-D school-based assessment (reporting in the same scale).

- To explore various approaches and evaluate options to arrive at a recommendation for selecting and assessing a nationally representative sample of non-institutionalised 14-16 year-olds who are out of school (as defined above) in future cycles of PISA, for countries choosing this option.

- To yield sufficient data to inform the countries' consideration of the policy implications for the out-of-school population.

With regard to the third objective, these policy implications might include, for example:

- meeting the challenge of achieving basic skills for all

- reducing grade repetition and school dropout

- eliminating barriers to schooling

- strengthening the evaluation and assessment system for schooling

- bringing dropouts back into schooling or skills training.

To accomplish these three objectives, PISA-D has developed a range of sampling procedures to identify and reach out-of-school youth, all of which were implemented in the participating countries by a national sampling manager responsible for all sampling-related activities, including design, selection and sample monitoring. This position has been created specifically for the PISA-D out-of-school assessment and needed to be filled by a specialist in each participating country with household-survey experience and knowledge of sampling hard-to-reach populations.

The PISA-D core sampling plan approach relied mostly on a probability sample, where all sample units have a non-zero probability of selection. PISA-D randomly selected households, and trained interviewers go to each household to apply a screener questionnaire to confirm if there were any eligible youth. If so, the interviewer then contacted the youth to conduct the interview or, if the youth was unavailable, set up an appointment to conduct the interview later. This procedure was complemented with a non-probability sample with referrals from schools or interviewed youth, or household members in households with no eligible youth.

The field trial included both probability- and non-probability-based samples. In the probability sample, which is more expensive, each unit has a known probability of belonging to the sample, whereas in the second case, there is no specific probability of an individual being selected. In the non-probability sample techniques are used for purposive selection of low-cost samples. These techniques include: a) link-tracing through households, for example by being referred to a household with out-of-school youth by another household; b) referrals to out-of-school youth from school teachers and students; c) use of school administrative records of students at or below grade 6; and d) special operations, for example used for reaching street children. The use of these sampling methods in the countries is indicated in the table below.

Table 5.2. Sampling methods used in the out-of-school assessment field trial

Country	Probability	Non-probability			
		Link tracing through households	School frame approach for OOS	School frame approach for <7th grade	Special operation for street children
Guatemala	X			X	
Honduras	X	X	X	X	
Panama	X	X	X	X	
Paraguay	X	X		X	X
Senegal	X	X			
Zambia	X	X			X

Note: OOS = out-of-school youth

5.3.2. Survey operations

The PISA-D out-of-school youth assessment is conducted on tablet computers and takes place at the youth's home. Participating countries test the assessment software that has been developed for the pilot and then train interviewers in the use of this software. The interviewer poses a 30-minute series of background questions to the youth covering topics such as their school and learning experiences, well-being and home life. The youth then uses the tablet to complete a self-administered 50-minute test composed of about 38 reading and mathematics questions. Parents (or the most knowledgeable person about the youth) also answer a paper-based questionnaire about the youth's background and

childhood experiences, and the interviewer completes a questionnaire about the location and characteristics of the youth's household.

5.4. Findings from the PISA-D field trial of the assessment of out-of-school youth

The PISA-D tests and contextual questionnaires for out-of-school youth were piloted in six countries: Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia. The field trial for the out-of-school assessment instruments took place from April to September 2017. More than 32 000 households in the participating countries were screened to determine the eligibility of youth in those settings for the out-of-school assessment, yielding about 2 300 completed cases. This represents a hit rate (number of cases completed per number of youth screened for eligibility) of around 7% in the probability-sample component. Another 1 500 cases were completed by eligible youth who were identified using a methodology that was not based on probability sampling (no response rate was computed for this component). Based on the results of the field trial, the instruments and survey operations were modified where necessary in preparation for the Main Survey.

Findings from the analysis of the PISA-D field trial data for the assessment of out-of-school youth are organised around four major field trial goals. Two of the goals are related to sampling and survey operations and instrumentation and are summarised in the table below.

Table 5.3. Findings from the field trial of the PISA-D out-of-school assessment

Field trial goals	Findings
1. Evaluate the sample design and selection, and survey operations procedures	The field trial was successful for the survey operations procedures in terms of efficiency and accuracy of data collection processes, assessing response rates for various subpopulations of interest, and efficiency and accuracy of data processing (including recoding) and data submission.
2. Provide information about measures of the quality of the survey instruments	The sample sizes across all participating countries were smaller than planned but were adequate for finalising the Main Survey instruments and design. The results of the field trial are a valuable resource to help guide the adequacy of scoring procedures, the quality of translation and adaptation, and the scale and analytical procedures for the Main Survey.
3. Evaluate scaling and psychometric item characteristics	The PISA-D field trial data provided valuable information on how well PISA items function, as well as items that came from other surveys and their appropriateness to the PISA-D out-of-school population. However, a lack of sufficient data at the item level across countries prevented the full set of FT analyses to be completed as planned. The analyses intended for the FT were postponed to the MS phase and contractors noted there was uncertainty about the appropriateness of the items for the Strand C population that would not be fully understood until after the MS analyses are completed.
4. Assess the use of the computer-delivery platform and the case management system	The field trial was a success in the following tasks related to an interview-based study: i) the functioning of the cognitive portion of the delivery platform was optimal in the collection of responses and automatic scoring; ii) the case management system worked correctly, particularly for the flow of questions and efficiency of the system in capturing information; iii) it was possible to evaluate the accuracy of the interviewer's instructions; iv) the case management system was effective during the interviews; and v) the system for assigning cases to interviewers, storing case files, and managing reports at the national level worked correctly.

The PISA technical standards were elaborated to account for the new features that the out-of-school assessment brings to PISA. The results of the field trial confirm that the participating countries carried out their survey operations in accordance with the newly elaborated technical standards, but with some difficulties. The results of the field trial are a valuable contribution towards the design of the Main Survey, and for accuracy of the data collection and submission practices in the out-of-school component. Based on the field trial results and experiences from countries, the international contractors also proposed to modify the main survey assessment design to accommodate a smaller sample size yield, while maintaining minimum yields necessary to evaluate the assessment items. These modifications will be reflected in the final version of the technical standards.

All findings from the PISA-D field trial, including lessons learnt, were documented for reference in preparing for the main data collection. While the results of the field trial were positive, each country faced its own challenges in conducting the out-of-school assessment, such as completing all field trial tasks on time. The field trial helps countries better plan for the Main Survey and anticipate challenges. The OECD and its contractors provided countries with tailored support to reinforce their capacity and create the conditions in each country that are needed to successfully implement the assessment.

5.5. Out-of-school sampling plan for the Main Survey

The goal of the Main Survey, which took place during the second half of 2018 and into 2019, was to achieve a minimum sample size of 1 600 in each participating country, with at least 1 300 youth completing the main assessment items. It was envisaged that this should result in an average of 650 responses per main assessment item. The 1 600 is split into at least 1 200 from a representative sample, and at most 400 from a limited representativeness sample. The representative sample adequately represents the target population, and consists of at least 600 cases from probability-selected households. As an option, some of the remaining cases could come from referrals to out-of-school youth in the neighbourhood from the probability-selected households.

A key aim was to maximise the sample size that comes from the probability sampled households, but at the same time allow additional cases to be completed through less expensive means to guarantee the achievement of sample sizes for the Main Survey. Some countries – particularly those with larger populations of out-of-school youth – have less difficulty in obtaining a representative sample and may not use the limited representative sample, or use less of it. Countries with small populations of out-of-school youth, or where these are located in locations difficult to access, will likely use all the limited representative sample that is permitted.

5.6. Out-of-school assessment data-collection operations for the Main Survey

The PISA-D out-of-school youth assessment was conducted through in-person interviews on tablet computers, and took place at the young person's home, in most cases. In a session lasting around 30 minutes, the interviewer posed a series of background questions to the 14-16 year-old respondent, covering topics such as the youth's school and learning experiences, well-being and home life (see Chapters 3 and 4 for more details). The interviewer recorded the responses on the tablet.

The test begins with a core module composed of five reading and five mathematics items. Depending upon the responses to these, the youth was guided automatically to either the full test or an assessment of reading components designed to paint a more nuanced picture of low performance (see Figure 5.1). Parents (or the person most knowledgeable about the young person) also answered a paper-based questionnaire about the youth's background and childhood experiences. The interviewer completed a short household-observation module on the tablet that included questions about the location of the household, aspects of the neighbourhood, and some characteristics of the dwelling.

Figure 5.1. PISA-D out-of-school main survey assessment design
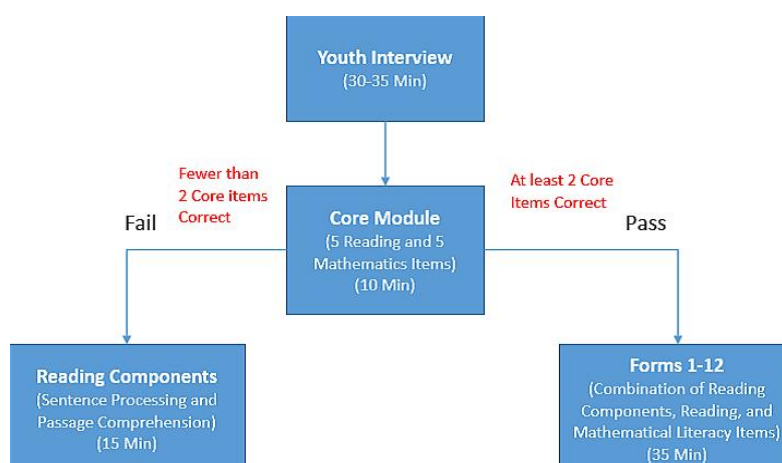
As noted earlier in this report, each participating country appointed a National Survey Operations Manager to take responsibility for all survey operations, including interviewer recruitment and training, and field operations management. The manager, supported by international contractors was responsible for ensuring that the PISA-D technical standards for the out-of-school component were followed at every stage.

5.6.1. Data collection and survey operations

The OECD's international contractors developed interviewer training materials for countries to adapt and translate as needed. In each country, qualified interviewers attended an in-person training, which included general interviewing techniques and project-specific training on the concepts, instruments and procedures related to PISA-D.

The in-person training was designed to maximise trainees' involvement and participation in the exercise, and to provide opportunities for supervisory staff to observe and evaluate trainee performance. The contractors also provided support to countries during the data collection, including through some country visits.

Countries submitted sample monitoring quality-control forms to the international contractor to help identify shortfalls in the sample, problems in achieving the desired response rate, and the potential for non-response bias in the collected sample.

5.7. Implementation of the out-of-school youth Main Survey

The PISA-D tests and contextual questionnaires for out-of-school youth were administered in five countries: Guatemala, Honduras, Panama, Paraguay and Senegal. The data collection for the out-of-school assessment instruments took place from June 2018 to January 2019. About 90 000 households in the participating countries were screened to determine the eligibility of youth in those settings for the out-of-school assessment, yielding a little under 7 500 completed cases. This represents a hit rate (number of cases completed per number of dwelling units screened for eligibility) of around 8% in the probability-sample component. Another 900 cases were completed by eligible youth who were identified using a methodology that was not based on probability sampling (no response rate was computed for this component). The instruments and survey operations had been modified where necessary for the Main Survey, based on the results of the field trial.

The number of completed cases achieved in each of the participating countries compared to the target completed cases is presented in Table 5.4 below. As will be noted, in two countries the actual number of completed cases is much lower than the target number owing to a lower than expected hit rate – in Honduras and Paraguay, out-of-school youth were twice as hard to identify, approach and interview than they were in the other countries. In the case of Guatemala, this country only narrowly missed the target number of completed cases, whilst Panama and Senegal exceeded the target number. Both Honduras and Panama had probability-based samples selected only in their rural areas. Each had small limited-representative samples in urban areas. Expected hit rates were checked against actual hit rates. In Panama, hit rates were much higher than expected in their rural and indigenous areas, and therefore some caution should be given regarding the representativeness of the sample. Building upon the existing quality control checks will be important in future cycles to ensure quality results under extraordinary challenge of reaching the out-of-school youth.

The use of probability-based link-tracing (asking probability selected households for referrals to eligible youth) is promising in terms of reduced cost, while retaining similar bias and variance. However, some limitations include that there is evidence that people do not like to talk about other people. For example, in the field trial, the approach did not work in Honduras urban areas, but worked okay in rural areas. Also, the home office needs to de-duplicate the referrals prior to assigning to interviewers. Lastly, there is loose control of Primary Sampling Unit boundaries; referred to those "in neighborhood".

For the main survey, all but Panama elected to use a mini-census approach, which involves screening all households in selected areas, while also logging disposition codes of those that are not successfully completed. The mini-census approach was preferred as a more straightforward approach when dwelling unit sampling frames were not available, when they had small-sized primary sampling units (around 50 dwelling units), and countries had outdated maps. Because some hit-rates in certain areas of certain countries were somewhat different than expected, improved quality control procedures are recommended to verify correct implementation of the mini-census approach.

An improvement toward better use of the probability-based linking tracing approach, as an optional alternative to referrals, is that households could be offered to use a coupon to allow the probability household member to hand to their in-scope acquaintance. The acquaintance can decide to call the home office to be interviewed (like Respondent Driven Sampling).

Table 5.4. Actual completed cases versus target completed cases

Country	Actual completes				Target completes			
	Representative sample		Limited representative sample	Total	Representative sample		Limited representative sample	Total
	Probability	Link-tracing			Probability	Link-tracing		
Guatemala	1 041	NA	482	1 523	1 200	NA	400	1 600
Honduras	1 070	NA	117	1 187	1 200	NA	400	1 600
Panama	1 902	81	66	1 908	1 000	200	400	1 600
Paraguay	669	NA	156	825	1 460	NA	140	1 600
Senegal	2 023	NA	NA	2 023	2 000	NA	NA	2 000

5.7.1. *Quality monitoring, data submission and data processing*

The five participating countries that attempted the Main Survey each submitted completed quality-control monitoring forms to the international contractor prior to, during and after data collection. These forms included information about production, field staff, quality control, and systems issues. Participating countries and international contractors engaged in monthly discussions about the data collection during the Main Survey data collection period. These discussions served to keep the contractor informed on the progress of various operational aspects of field work and provided countries with the opportunity to obtain guidance on specific challenges faced, vis-à-vis data collection and related quality control measures.

At the end of the data collection period, each PISA-D participant country submitted its data in a single database. The scheduled submissions ensured that the international contractors had adequate time to review the data and provide feedback to countries. The data were processed and analysed in accordance with the timeline outlined below in Table 5.5. It is important to note that this timeline was lengthened to December 2020 in the light of the Covid-19 pandemic that became a global crisis from February 2020.

Table 5.5. PISA-D out-of-school Main Survey data processing, analysis and reporting timeline

Timeline	Data processing	Data analysis and reporting
1 June – 31 July 2019	Cleaned national data files to countries for review and comment. Contractors finalize data cleaning based on feedback from contractors and countries	Contractors and OECD begin analysis of data, complete scaling of background questionnaires data and develop proxy values to facilitate initial analysis
1 August – 8 November 2019	National data files with weights delivered to National Centres – countries provide feedback on weights and these are then finalised Contractors generate proficiency levels	Analysis and national report templates updated to incorporate weights – lead analysts in participating countries draft reports
December 2019	International Database 1 Release and national data files/Suppression Requests processed	Panama launches its PISA 2018 national report, incorporating analysis of its PISA-D out-of-school data
1 January – 30 November 2020	International database finalised and shared with OECD and participating countries Technical Report prepared by contractors	OECD PISA analysts and contractors work with country teams, in particular the lead analysts from each country, to analyse the out-of-school data and to prepare national reports
1 – 3 December 2020	International Database 1 published Technical Report on Strand C published	Four countries launch their national reports, incorporating the analysis of out-of-school data. OECD launches PISA-D Results-in-Focus (out-of-school), the International Database 1, Technical Report and final version of Project Completion Report

5.8. Findings from the PISA-D Main Survey of the out-of-school assessment

The main survey data collection was subject to a strict adjudication process, particularly for the sampling and translation/adaptation parts of the implementation. The way the sample was selected and the instruments were localised in each country was an integral part of the evidence supporting the construct validity of the intended uses of the assessment. Overall, the PISA-D adjudication review suggests that the PISA-D out-of-school assessment has been implemented in accordance with the technical standards that were developed for this pilot and that the isolated deviations from these standards did not affect the quality of the data products. Overall, the PISA-D Strand C quality control programme for survey operations met the intended goals. During the Main Survey data collection:

- country compliance was satisfactory

- the OECD and the contractors were kept informed about the progress of data collection

- countries were supported by having their questions answered by the OECD and its contractors, and any areas of concern were pointed out promptly throughout the critical months before and during data collection

- the programme experience will serve countries and the OECD as they plan for future cycles of PISA-D Strand C.

During the analysis phase of the project, analysts took an in-depth look at the results, including the functionality of all test items, and confirmed that the instruments measure what they purport to measure against the *PISA-D Assessment and Analytical Framework* (OECD, 2018[5]). Indeed, the analysis validated the assumptions that shaped the framework. As a result, participating countries can be confident that the assessment results will provide relevant data that can inform decisions concerning national policies.

5.8.1. International comparisons

The project achieved its primary goal of scaling PISA-D Strand C to provide a reliable and valid link to the PISA scale (i.e. the scores from both assessments can be located on a comparable scale) by linking PISA-D Strand C to PISA-D Strand A/B through fixed item parameter linking. This achievement was facilitated by having all of the items in PISA-D Strand C, with the exception of one item (a reading component item), selected from PISA-D Strand A/B and the scoring rules were also the same as those applied in PISA-D Strand A/B. PISA-D was therefore successful in making the assessment instruments for out-of-school youth relevant for this sub-population in LMIC while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA. These variables and factors include performance on the test (e.g. Table 5.6), educational attainment, health and well-being, attitudes towards school when attended and learning, family and community support.

Table 5.6. Proficiency in reading amongst in-school and out-of-school youth

	Percentage of 15-year-olds covered by in-school assessment	In-school students performing at Level 2 or above	Out-of-school youth			Percentage of 15-year-ods performing at or above level 2 (weighted average)	Interquartile range of proficiency level	
			Performing at or above Level 1b	Performing at or above Level 1a	Performing at or above Level 2		Out-of-school	In-school
	%	%	%	%	%	%		
Guatemala	47.5	29.9	42.1	7.7	0.4	14.4	1C-1B	1B-2
Honduras	41.4	29.7	65.2	21.3	2.2	13.6	1C-1B	1B-2
Panama	53.5	35.7	60.2	25.3	4.1	21.0	1C-1A	1B-2
Paraguay	m	32.2	37.7	6.0	1.0	-	1C-1B	1B-2
Senegal	29.0	8.7	33.9	3.0	0.0	2.5	1C-1B	1C-1A
Out-of-school average	42.9	27.2	47.8	12.7	1.5	12.9	1C-1B	1B-2
OECD average	88.0	76.1	-	-	-	-	-	2-4

Source: PISA 2018 Database and PISA for Development Database, Strands A, B and C.

5.8.2. Policy insights

PISA-D assessment results provide countries with a solid database that can help them refine policy priorities and set new goals or targets to improve their education systems. The data collected have a lot to say about the allocation of resources and its implications for equity. With reliable data on gaps in access and differences in outcomes between groups of children and young people, countries can determine whether poor and marginalised populations are given equal opportunities to succeed at school and beyond. The challenge for countries is to maintain a focus on these goals or targets, and to track progress towards them by participating in future cycles of PISA and other relevant studies.

A key rationale for the PISA-D out-of-school assessment pilot is that measures of school achievement through administration of tests refer to students - not the whole population – and this poses a problem for assessing efficiency and human capital, especially in low-and-middle-income-countries. With regard to the efficiency of an educational system, it is important to see test scores in the context of participation rates and the extent to which the dropout rate in a country has been reduced between any two measurement points. With regard to evaluating human capital – the whole population – it is essential to know

the human capital of school dropouts and those that never enrolled. In addition, dropout and never enrolling is very negatively correlated with socio-economic status. An assessment in a country with low education system coverage that ignored the out-of-school is at risk of perversely encouraging policies of exclusion.

While it is possible to estimate test scores for the whole population (i.e. taking into account dropouts and those who never enrolled) by putting bounds on unobserved scores, this is effectively guess-work and is carried out under weak assumptions. There is no substitute for assessing the skills of the whole population as has been done in the PISA-D countries.

5.8.3. Next steps for scaling up the PISA-D pilot

The OECD is exploring three options for taking the PISA-D out-of-school assessment pilot forward:

- A *stand-alone* international option as part of a future PISA cycle (the PISA-D Strand C test with results linked to the PISA scale).

- An international option as part of a future PISA cycle *linked to a household survey* (the PISA-D test with results linked to the PISA scale).

- A *PISA mini-test (a 15-20 minutes test designed from the PISA-D Strand C instrument) integrated with a multi-topic household survey* designed solely to discriminate whether respondent is above or below 406 points on reading and 419 points on mathematics – the SDG 4 benchmarks for minimum levels of proficiency – may be part of a future PISA cycle or a completely separate study.

With regard to the first and second of these options, the OECD is interested in working with other PISA participants, like the PISA-D participants, that have large proportions of youth that are outside PISA's target grades. While several countries have expanded access to education for their 15-year-olds in recent years, there are still several PISA 2018 participants where less than 80% of the population of 15-year-olds were covered by the PISA sample (meaning that they were enrolled in school, in grade 7 or above), see Figure 5.2 below.

Figure 5.2. Coverage of the national 15-year-old population in PISA 2018 (Coverage Index 3)
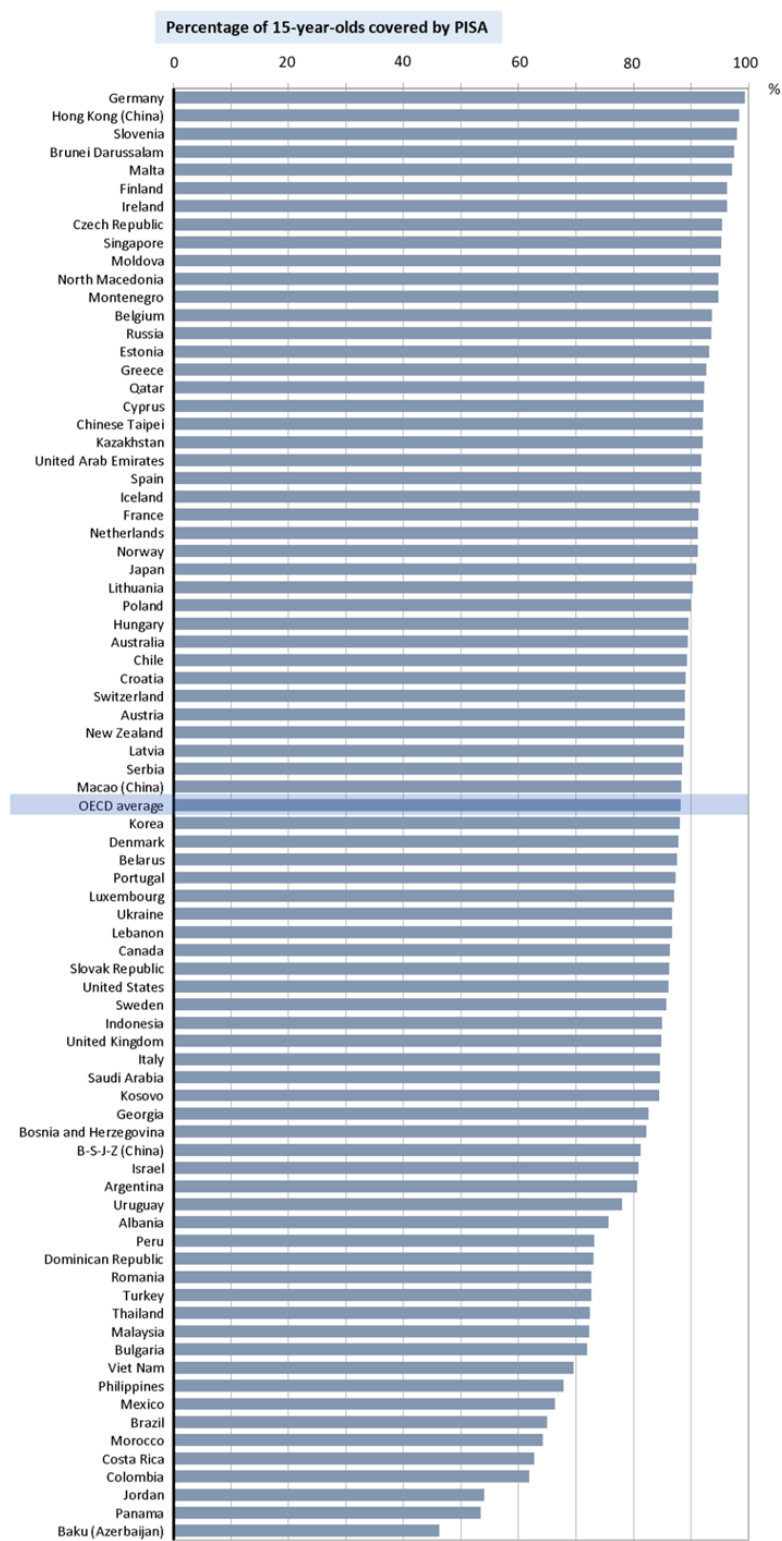
With regard to the third option, the OECD is working with the World Bank, UNESCO and UNICEF in particular to develop a concrete roadmap for improving learning data collection in household surveys through integrating learning assessment modules with these, including the proposed PISA mini-test. In addition to a roadmap, a guidelines document will also be developed for how to capture individual abilities to meet certain minimum thresholds ('foundational skills'). The OECD and its partners are hoping to identify 2-3 countries to pilot the approach(es) for integrating learning assessment modules in household surveys over the course of 2021 and 2022. This would then be taken to scale as part of efforts to close learning assessment data gaps for SDG 4, especially in low-and-middle-income-countries.

5.9. Achievement of the output

This output of the project has been achieved within the scope of the project budget. An analytical framework and methodological approach for including out-of-school 15-year-olds in assessments has been developed and piloted. As noted above, these approaches are now being scaled up both within future cycles of PISA and as part of household surveys that may or may not be implemented in conjunction with PISA cycles. Approaches to out-of-school children have been developed and trialled according to the plan agreed, but the schedule has been delayed by 18 months. As noted in Chapter 2, the causes of these delays were related to the delivery of the tablet computers to the countries (3 months), a longer than expected period for the Main Survey data collection (6 months) and the impact of the global Covid-19 pandemic (9 months).

6. Capacity building of participating countries in assessment and analysis

One of the main aims of PISA-D was to help to build the capacity of participating countries to conduct large-scale learning assessments, and analyse and use the results to support national policies and evidence-based decision making. One of the most significant achievements of PISA-D is the active engagement of the participating countries, not just in the implementation of PISA, but also in the analysis and use of the data generated by the assessment.

Of the nine countries (Bhutan, Cambodia, Ecuador, Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia) that partnered with the OECD to develop and test the enhanced PISA instruments, Panama is the only country to have participated in PISA previously. However, all the countries have experience with regional or international assessments and conduct national student assessments.

6.1. Preparing countries to participate in PISA-D and to make use of the results

PISA-D countries were prepared for their participation through a three-stage process that started with an analysis of their capacity to implement PISA and make use of PISA data, and included planning to strengthen that capacity. Countries were supported by the OECD and its contractors at each stage of the PISA cycle:

Capacity needs analysis ➡ Capacity-building plan ➡ PISA implementation plan ➡ PISA implementation support, including technical assistance in support of survey operations

This process was designed to help LMIC overcome two potential barriers to their participation in PISA: a lack of capacity to implement the assessment and a lack of experience in using PISA data and results. To overcome the latter obstacle, in PISA-D the OECD and its contractors provided training and assistance in data analysis, the interpretation of PISA results, report writing and communication through the following process:

Preparing stakeholders for PISA results ➡ Analysis of PISA results and country report ➡ Dissemination of results and evidence-based policy dialogue ➡ Better policies for improved learning outcomes

The capacities strengthened through PISA-D are relevant to countries' management of their own national assessments and other large-scale international or regional assessments in which they might participate. As noted earlier in this report, all of the National Centres responsible for implementing PISA were government institutions that were also responsible for other learning assessments in the countries, international, regional and national.

6.1.1. Getting ready to participate in PISA

A three-stage process was developed to prepare countries for PISA-D participation:

1. Capacity Needs Analysis (CNA): ensures there is a solid foundational capacity for implementing the project and identifies areas of potential growth for the country

2. Capacity Building Plan (CBP): addresses identified capacity needs and enhances the enabling environment for PISA, particularly the use of assessment results for national policy dialogue and evidence-based decision making

3. Project Implementation Plan (PIP): describes the actions to be carried out by the specific entities and agents that are named and commissioned for implementation by the authorities of the participating country, together with the necessary resources.

The CNA framework that was developed and applied to PISA-D countries is composed of three dimensions: the enabling environment; the organisation; and the individual. The World Bank's Systems Approach for Better Education Results (SABER) Assessment rubrics and PISA standards were combined as the benchmarks for the assessment of capacity in the participating countries. The analysis was conducted by consultants contracted by the OECD to visit each PISA-D country and work with relevant national staff to determine the capacities that required development in order to implement PISA in accordance with the technical standards. The findings of the CNAs for all countries were positive with most of the capacity assets necessary for project implementation either emerging or established.

The CBP template was designed to prioritise the immediate capacity requirements within the PISA-D project cycle and to then address additional capacity building items requested by the countries. On the basis of the CAN findings, countries moved into the phase of designing their CBPs with the support of the OECD consultants. The CBPs were then reflected in each country's PIP as well as in the terms of reference for the work of the PISA-D contractors.

The PIP succinctly describes the actions to be carried out by the specific entities and agents that are named and commissioned for implementation by the authorities of the participating countries. The PIP also collects information about policy priorities and expectations, communication strategy, risk management strategy, evaluation and monitoring strategy, and perhaps most importantly a fully costed project budget.

All CNAs and CBPs are available on the OECD website[7]. The PIPs are internal working documents for the OECD and PISA-D national team's use only.

Participating countries each nominated a National Centre and nominated a National Project Manager to ensure appropriate infrastructure and resources were in place to implement the assessment in accordance with the PISA technical standards. As already noted, these National Centres were all government institutions that were responsible for other large-scale assessments, including national assessments, and the National Project Managers were responsible for other learning assessments, not just PISA. These arrangements ensured that capacity building from PISA flowed directly into the management and implementation of all the large-scale learning assessments that the countries were engaged in.

---

[7] www.oecd.org/pisa/pisa-for-development/pisa-for-development-documentation.htm

Feedback from PISA-D countries on this part of the project indicates that they appreciated the needs analyses conducted at the beginning of the project as both a useful assessment of the national team's strengths and weaknesses and as a benchmark for assessing how far they have come in capacity building since the project began.

As preparing countries for their participation in PISA-D and their use of the data through this three-stage process worked well, it was made available to new countries participating in PISA 2018 with Panama and Ukraine being the main beneficiaries, and it has continued being used in PISA 2022. In PISA 2022, four countries (El Salvador, India, Mongolia and Uzbekistan) are benefiting from these processes and activities.

6.2. Capacity building

To strengthen each country's participation in PISA-D, the OECD sought to bolster national capacity for managing large-scale student learning assessment and using the results to support policy dialogue and decision making. The OECD provided participating countries training on a variety of topics identified in the reports of the capacity needs analysis, including framework and item development, sampling, translation/adaptation of survey instruments, data management, coding of students' responses, data analysis and reporting.

6.2.1 International National Project Managers' Meetings (NPMs)

Capacity building was conducted primarily through PISA-D National Project Manager meetings and other international training events. As described in Chapter 2, ten of these meetings were scheduled during the 2015-18 period and were attended by all country teams. The training was delivered by OECD officials and OECD contractors. Country representatives who attended the meetings were responsible for disseminating the training to the rest of the staff at their national centres. Country representatives at these meetings included the NPMs, data managers, assessment specialists, coders and lead analysts – as noted earlier in this report, all of these country representatives are also responsible for other learning assessments in their countries, not just PISA. The OECD and its contractors complemented the international training events with weekly calls, scheduled teleconferences, webinars and country visits. PISA-D development partners provided additional support, such as the Inter-American Development Bank, which conducted regional training workshops in Spanish for PISA-D countries in Latin America. Representatives from other Latin American PISA-participating countries and from Latin American countries that do not (yet) participate in PISA also participated in these training workshops.

NPM 1

The first international/NPM meeting from 28 September to 2 October 2015 in Quito, Ecuador. The participating countries demonstrated a high level of engagement in the technical development work and the first international meeting, providing useful comments and endorsing the draft frameworks for the assessment and the initial selection of items for both the tests and the questionnaires. This ensured that the project stayed on track and in accordance with the agreed implementation schedule. The first international meeting also greatly benefited from contributions by the following peer learning countries: Chile, Korea, Mexico, Panama, Peru and Uruguay.

NPM 2

The second international/NPM meeting was held in January 2016 at the offices of Westat in Rockville, US. With the frameworks and instruments for the field trial agreed, the second meeting concentrated on preparing the participating countries for the tasks of adapting, translating and printing the instruments and managing the survey operations. The meeting also included several sessions on Strand C, particularly to agree on sampling issues and the plans for developing instruments. As a result of the meeting Strands A and B remained on track and there was significant momentum to both efforts. Strand C also started to take shape and contractors and countries filled the gaps sensibly, productively and in a good team spirit. There were challenges to all efforts but the good working relationships forged between the countries and the contractors and among the countries themselves and with their PISA peers and helped ensure that countries and contractors met all challenges successfully. Peer-to peer learning contributions were made at the meeting by United States and Canada and were well received.

NPM 3

The third international/NPM meeting took place from 4-8 April 2016 in Asunción, Paraguay and covered the sampling plan for the PISA-D in-school assessment, including detailed discussion with each participating country to finalise individual national sampling plans. The meeting also covered the survey operations more generally. In addition, the test administration procedures were set out in preparation for the fourth NPM meeting.

NPM 4

The fourth international/NPM meeting took place from 4-8 July 2016 in Livingstone, Zambia and covered coding reading, mathematics, science and occupations; survey operations for the field trial; data management training; and ICT training for the out-of-school assessment. With regard to the ICT training, Microsoft's head of education for Africa, Phil Oduor, brought six tablets to the meeting and supported ETS's delivery of the training. Microsoft and its partner, Positivo, also confirmed they would provide support to the countries as they receive the tablets, up-load the software and train the interviewers. At this meeting, the participating countries were prepared for conducting the field trial, the coding training was delivered, and the arrangements for data management and the preparation of the tablets for the out-of-school assessment were reviewed.

NPM 4A

Following the fourth international/NPM meeting, a Strand C-specific international/NPM meeting, referred to as 4A, took place from 1-4 November 2016 in Madrid, Spain. The meeting focused on Strand C and covered interviewer training, using interactive role plays with the tablets and interview software; survey operations for the field trial, including sampling updates and the tablet-based delivery system; and data management training.

NPM 5

The fifth international/NPM meeting took place from 22-25 May 2017 in Siem Reap, Cambodia and included a review of the Strands A and B Main Survey design, Technical Standards, instruments and questionnaires for the PISA-D in-school assessment. Procedures for changes to the instruments and questionnaires were confirmed, and procedures for sampling were described. Survey operations were planned to take the project through the Main Survey of Strands A and B as well as for the Field Trial phase of Strand C.

In addition, the test administration procedures were set out in preparation for the next international/NPM meeting

NPM 6

The sixth international/NPM meeting took place from 24-28 July 2017 in Princeton, New Jersey, United States and focused primarily on Strands A and B, covering a status update of translations, adaptations and preparation of the Main Survey booklets; student sampling for the Main Survey; training on coding of open ended questions, data management and coding of occupations; quality checks; and survey operations, preparing for and conducting the Main Survey. The meeting also included a sampling workshop for Strand C. The meeting benefited from peer learning experiences with Canada and the United States, represented by Pierre Brochu and Dave Kastberg, respectively.

NPM 7

The seventh international/NPM meeting took place from 7-11 May 2018 in Saly, Senegal and focused on the analysis of the MS data for Strands A and B and lessons from the field trial as well as the arrangements for the main data collection for Strand C. The steps regarding the analysis and checks of the Strands A and B Main Survey data files were reviewed and the countries confirmed their agreement. The Main Survey instruments and questionnaires and procedures for Strand C were reviewed and the countries confirmed their agreement. The analytic techniques and tools for the analysis of the Strands A and B data were reviewed and agreed and relevant training was provided to the participants. The adjudication process for the Main Survey of Strands A and B were reviewed and key steps identified for noting by the countries. There was a peer learning session with Korea, a session on the structure of the PISA-D national reports, a session on the lead analysts programme, a session on engagement and communication, and a session on the process for data adjudication.

NPM 8

The eighth international/NPM meeting took place from 23-27 July 2018 in Antigua, Guatemala and focused on completing the national reports for Strands A and B and advancing the arrangements for the main data collection for Strand C. The draft national reports for Strands A and B were reviewed and agreed by the countries and the steps regarding the finalising of these were reviewed and the countries confirmed their agreement. Training and peer learning from Dominican Republic and Mexico was provided on policy implications of PISA results, disseminating PISA results, especially bad results, communicating with the media and data visualisation. The arrangements for the launches of the national reports in the week beginning 10 December 2018 were reviewed and agreed with the countries. Training and hands on experience was provided on using the Data Explorer for the analysis of the Strands A and B data and the participants were made aware of the potential of the Data Explorer for maximising the utility of the data, disseminating the results and encouraging secondary research. The next steps for the Main Survey data collection for Strand C were reviewed and the countries confirmed their agreement – training was provided in data workflow and collection, sampling and survey operations, case management and data management. Next steps were set out and agreed with the countries, including the processing of their participation in PISA 2022.

6.2.2 Other capacity building activities

In addition to the international NPM meetings, the following capacity building activities took place as part of PISA-D:

- June 2016: IDB-supported training on assessment frameworks and item preparation hosted in Ecuador for Latin American countries (40 participants representing Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay and Peru)

- August 2016: ETS/Westat/OECD capacity-building visit to Senegal focused on the out-of-school assessment

- September 2016: IDB-supported training on Item Response Theory hosted in Guatemala for all the Latin American countries (36 participants representing Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay and Peru)

- June 2017: IDB-supported training on secondary analysis hosted in Honduras for all the Latin American countries (36 participants representing Ecuador, El Salvador, Guatemala, Honduras, Panama and Paraguay).

- September 2017-December 2018: Implementation by the OECD and its contractors of the programme for the Lead Analysts. During the OECD residency portion of the programme in 2018, Lead Analysts learned from numerous presentations including by UNESCO representatives on the Global Education Monitoring Report for 2018 and by World representatives on the World Development Report for 2018.

- April – December 2019: Participation by Panama in the OECD's PISA 2018 Lead Analysts Programme. During the OECD residency portion of the programme in 2019, Panama's Lead Analyst learned from numerous presentations including by UNESCO representatives on the Global Education Monitoring Report for 2019

- January-September 2020: virtual discussions and learning sessions provided by the OECD for the Lead Analysts to support their analysis and reporting on the Strand C data

6.3. Analysis of the PISA-D data and production of the PISA-D national reports

With the support of the OECD, each PISA-D country formed an analysis task force and national report working group, and nominated one analyst to lead the work on data analysis and reporting. The analysis task forces and working groups in each country were made up of staff from the National Centre and Ministry of Education, from departments that had a direct engagement in the results of the assessment. The OECD worked directly with the analysis task force, especially the lead analyst from each country and implemented a lead analysts' programme from September 2017 through December 2018 for analysis and reporting of the school-based assessment results, see Table 6.1 below.

Table 6.1. PISA-D lead analyst programme

Timeframe	Learning targets
September-December 2017	Introduction to PISA data and its uses for policy makers, educators, researchers; guided reading
January-May 2018	Preparing for analysis of Main Survey data: understanding the PISA design; limitations of PISA samples; understanding and interpreting PISA coverage indices; the analytical framework
	Drafting of introductory Chapter 1 of national report
May-July 2018	Analysis of data and drafting of Chapters 2-6 of national report: training in data-analysis techniques and methodologies, collaborative drafting with OECD (residence at OECD, Paris)
July-November 2018	Finalising the national reports and preparing for launches, including producing country-specific dissemination materials (PPTs, country notes, Tweets, videos, webpages, etc.)
December 2018	Publication and dissemination of national reports with final results from the school-based assessment

The first period of the residence at the OECD for the Lead Analysts (May) included training on PISA data analysis and reporting; the characteristics of international assessments and examples of research conducted using their initial PISA-D data files; how to handle international studies' databases and technical documents; how to replicate PISA results and answer research questions using relevant software (such as Mplus, R, STATA, SPSS or SAS); and practical experience on how to formulate, analyse, interpret and present results relevant to educational research using PISA-D data. In the second period of the residence (June to July) the Lead Analysts applied the skills acquired through the training to analysis of their own PISA-D data and, with the mentoring and support of the OECD's PISA-D manager, PISA Analysts and Experts interpreted their results and prepared a report on the basis of the OECD's national report template, but adapted to suit their national contexts and education priorities.

During the period August – November 2018, the Lead Analysts returned to their countries, consulted with the relevant authorities on their draft reports, and obtained the necessary clearances for the publication of the report by their respective Ministries of Education. This period also included the completion of any outstanding data analysis in light of the final release of the international database to the countries, review and finalisation of the national reports, the production of these and the preparation of associated communication products, all of which was supported remotely by the OECD and its contractors.

The aim of the Lead Analysts Programme was to help the Lead Analysts complete their analysis of PISA-D data and prepare a national report. Each report, containing six chapters, was based on a template prepared by the OECD and included analyses of the four key outcomes, known as "Prosperity Outcomes", for each stage of schooling and child development, and the "Foundations for Success" that underpin them (see Figure 6.1 below). During the writing of the national reports, this template was adapted by the Lead Analysts according to their countries' contexts and education priorities. The national reports were launched in the countries supported by the OECD and in conjunction with the launch of the OECD's PISA-D Results in Focus publication, the PISA-D Technical Report and the international PISA-D datasets in December 2018.

The Lead Analyst from Panama participated in the OECD's PISA 2018 Lead Analysts Programme which followed a similar course during 2019 to the PISA-D Lead Analysts Programme the previous year and included six other countries in addition to Panama, namely: Albania, Bosnia-Herzegovina, Indonesia, Moldova, Serbia and Ukraine. In addition, during 2020 the OECD and its contractors provided similar levels of support, virtually, to Lead Analysts in four countries (Guatemala, Honduras, Paraguay and Senegal) for the analysis and reporting on the Strand C data. The national reports on the Strand C

data, together with the OECD's international data products for this Strand of the project, were launched in December 2020.

Figure 6.1. PISA-D analytical framework
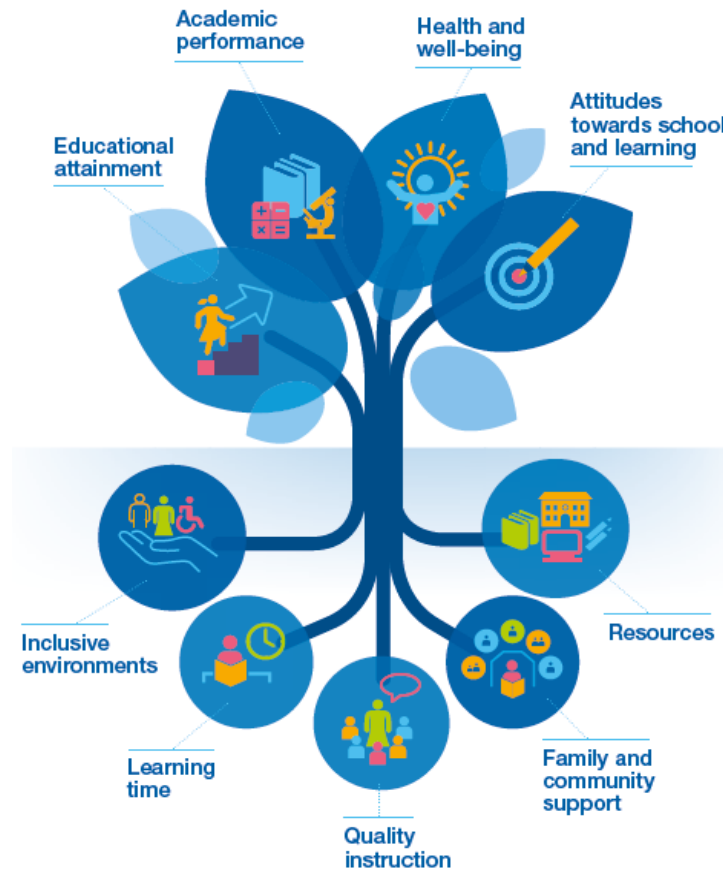
6.3.1 Preparing stakeholders for PISA results

The national reports are the culmination of an engagement and communication strategy that was implemented by each country from the start of the project. These strategies involved key stakeholders in each country in the survey and the discussion of the results and implications for policy. Stakeholders include pupils, parents, teachers, teacher unions, school principals, academia, civil society, media and central and local government.

When each country joined PISA-D, a launch event was held to mark their participation in the project and introduce PISA to stakeholders. From there, each country linked a communication focal point in its National Centre with the OECD, and the focal point led implementation of their national communication strategies in alignment with the PISA-D project's overall engagement and communication strategy. The OECD's strategy included the following activities:

- Key messaging to ensure people (including media) understand what PISA is and how PISA results can be used to improve learning, and change behaviour and perceptions where necessary.

- Tools to engage effectively with stakeholders (online effective communications presence, visibility tools, networking, meetings, workshops, communications circulated to countries participating in the project, peer-to-peer learning opportunities and an international seminar scheduled for the final year of the project).

- Continued feedback and analytics to demonstrate the impact of the project.

- Support for and alignment with the engagement and communication strategies of the participating countries in PISA-D.

The OECD produced a *PISA for Development Brief* series of over 30 issues, each two pages in length and describing a specific PISA-D topic. This series was intended to facilitate communication about the project for both the OECD and PISA-D countries, and is available online[8]. The OECD also maintained an up-to-date project website, which linked to participating countries' relevant websites. The website includes links to a brochure about the project as well as the *PISA-D Assessment and Analytical Framework*, both available in English, French and Spanish. In addition the OECD and its partners published blogs to promote the launches of national reports and other key documents produced through the project. These were some of the main communications tools produced by the OECD, and PISA-D countries produced their own tools including brochures, stakeholder-specific pamphlets, web pages, social media channels, journal articles, etc. The OECD reviewed these products and provided feedback to help countries strengthen their communications efforts.

Prior to and during the main survey data collection, the OECD helped prepare stakeholders in the participating countries to receive the results through various communications channels, including articles in newspapers and journals, pamphlets for teachers, social media campaigns, and TV and radio broadcasts. The OECD also participated in each country's national report launch event.

6.3.2 Dissemination of results and evidence-based policy dialogue for improved learning outcomes

The launches of the national reports and PISA-D data for the school-based assessment results took place in the PISA-D countries during the week of 10 December 2018 (Bhutan's launch took place in February 2019). The national reports and the PISA-D data enable stakeholders, especially policy makers, to gauge the level of knowledge and skills of students in their own country in comparison with those in other participating countries. If they wish, policy makers will be able to establish evidence-based benchmarks for improvement, such as in attainment and achievement or in closing gaps in education outcomes and opportunities between boys and girls or between urban and rural areas. The national reports, the PISA-D data, the launch events and subsequent discussions helped countries better understand the opportunities at hand and the challenges they face, and provided insights into how to help their students learn better, their teachers teach better and their school systems operate more effectively. All national reports are available online[9], as well as the PISA-D Database[10].

---

[8] www.oecd.org/pisa/pisa-for-development/pisafordevelopment-documentation-briefs.htm

[9] www.oecd.org/pisa/pisa-for-development/pisa-for-development-documentation.htm

[10] www.oecd.org/pisa/pisa-for-development/database/

The launches of the national reports were successful and were led in each case by the respective Ministers of Education and supported by the whole of government. Since the launches at the end of 2018, the participating countries report a high level of engagement in the national policy discussions that have followed. In Cambodia, for example, the Minister of Education has highlighted the PISA-D results as a factor in several policy decisions made during 2019.

The PISA-D analysis and reporting support model was successfully mainstreamed into PISA from the 2018 cycle onwards so that any participating country can sign up to collaborate with the OECD in analysing their data, interpreting their results, preparing a national report and disseminating the results to their populations. During 2019, seven countries (Albania, Bosnia-Herzegovina, Indonesia, Moldova, Panama, Serbia and Ukraine) loaned their Lead Analysts to the OECD for the period April to September to follow a similar programme to the PISA-D Lead Analysts Programme – these seven countries launched their PISA 2018 national reports written in collaboration with the OECD on 3rd December 2019 at the same time as the OECD launched its international report on PISA 2018. In the case of Panama, this country's national report also included the results of its Strand C survey.

As noted above, during 2020 the OECD and its contractors provided a high level of virtual support to Guatemala, Honduras, Paraguay and Senegal to help these four countries analyse their Strand C data and to produce national reports on this. These national reports were launched by the four countries on 1st December 2020.

6.4. PISA-D/PISA peer learning partnerships

An important innovation in PISA-D is the introduction of peer-learning partnerships between PISA-D countries and countries that already have experience with PISA. Peer learning is a feature of every international training event with experienced PISA participants sharing their lessons and experiences with the PISA-D country representatives. The following peer learning partnerships were established:

Table 6.2. PISA-D and PISA peer learning partnerships

Peers	Peer learners
Korea	Cambodia
Uruguay / Chile	Ecuador
Peru	Guatemala
Brazil	Honduras
Mexico / Colombia	Panama
Chile	Paraguay
Canada	Senegal

Some partnerships were stronger than others, and the strength of the partnership generally depended on the availability of resources for the peer learning activities. While the peer learning component of PISA-D was beneficial to most countries, it was difficult to organise a peer learning partnership for Zambia. The OECD discussed possible peer learning partnerships for this country from China, Hong Kong and Iceland during the course of the project. While all three of these PISA participants confirmed their willingness to provide peer learning support to Zambia, unfortunately, it was not been possible to identify the necessary funding to support their peer learning activities. Subsequently, the OECD and ETS proposed support for Zambia from experts in large-scale assessment that are based in

South Africa and to utilise the capacity building budgets allocated to the international contractors for this purpose.

The examples presented below are of formal peer learning activities/meetings/visits, but it should be noted that a lot of the peer learning in PISA-D also happened in day to day exchanges between countries via email or phone.

6.4.1. Examples of peer learning activities in 2016

- May 2016: Cambodia-Korea peer learning:

  o Extensive presentations from the OECD and representatives of Cambodia and Korea on the PISA-D project and the opportunities for greater participation in the assessment by LMIC

  o Extensive training and support from Korea's PISA team (KICE) for Cambodia covering all aspects of survey operations

- October 2016: Paraguay-Chile peer learning on coding and data entry held in Paraguay

- December 2016: First of three workshops in Senegal to build capacity for IRT, secondary analysis and item development led by a French-speaking consultant

6.4.2. Examples of peer learning activities in 2017

- March 2017: Colombia-Panama peer learning on sampling, producing the assessment instruments, timing and logistics of the assessment, analysis and reporting of results, and plausible values and oversampling held in Colombia

- March 2017: Paraguay-Chile peer learning held in Paraguay on Strand C technical aspects in view of instrument administration, in particular, preparation of the training material for the interviewers who will carry out the field trial of the out-of-school assessment; and on Strands A and B field trial results, expectations regarding what the results of the project can say, and preparing to present results with a national emphasis.

- May 2017: Seminar on assessment and curriculum organised by KICE (Korea) as part of the capacity development for Cambodia, which had recently reviewed its curriculum framework. The new framework outlines the knowledge, skills and attitudes that students are expected to develop and puts citizenship at the heart of the educational goal. The seminar aimed to help link the PISA-D project to the on-going discussion on the curriculum in Cambodia.

- August 2017: Panama received a peer learning visit from the Paraguay NPM (with financial support from IDB) to support the implementation of the Strand C Field Trial.

- September 2017: Paraguay received a peer learning visit from Chile (with financial support from IDB) to support their Main Survey implementation

6.4.3. Examples of peer learning activities in 2018

- May 2018: Peer-to-peer learning sessions were held at the IAG and NPM meetings held in Saly, Senegal. These sessions included representatives of France, Luxembourg and Korea presenting their countries' experiences with

communicating with stakeholders about PISA 2015, national data analysis, national report writing and result dissemination

- July 2018: A peer-to-peer learning session was held at the NPM meeting held in Antigua, Guatemala. This session included representatives of Dominican Republic and Mexico presenting their countries' experiences with communicating with stakeholders about PISA 2015, national data analysis, national report writing and result dissemination. The Dominican Republic representative also shared her country's experience with PISA's computer based assessment.

- August 2018: A peer-to-peer 4 day workshop for the Latin-American countries was held in Paraguay to help them prepare for the communication of results. Nineteen participants assisted from Ecuador, Guatemala, Panama and Paraguay, as well as representatives from Mexico, Chile, Peru, Dominican Republic and Costa Rica, and the representative from Uruguay who joined virtually. This activity had financial support from IDB.

In addition to the activities described above, the international contractors engaged with the countries on a continuous basis with weekly calls and scheduled tele-conferences and webinars. The contractors and the OECD closely monitored the countries' progress against the clearly defined tasks at each stage of survey implementation and intervened to support countries whenever problems or delays arise. ETS in particular facilitated capacity building inputs from experienced PISA specialists that speak French (Pierre Brochu), Spanish (Eugenio Gonzalez) and that are Africa-based (Sarah Howie) to ensure that the countries had all the support they need to implement the assessment successfully.

6.5. Country feedback on capacity building

At the annual PISA-D IAG meetings from 2016-18, each PISA-D country gave a presentation that covered the following: how many members of their teams have benefited from participation in an international NPM meeting or other capacity-building event since the last IAG; how they disseminated what was learnt at these meetings to other members of their team; the progress that has been made with regional capacity-building workshops; how the peer-to-peer learning partnerships are evolving; the aspect of the project that has been most helpful in building capacity thus far, and whether it has benefitted their national assessment in any way; and what they are most looking forward to in the next year for capacity building. They were also asked to comment on any important areas of capacity building that the project or other partnerships are not yet addressing or planning to address.

Countries largely praised the continuous support they received through every phase of the project. Countries also highlighted the pertinence of the training provided, including sampling, translation/adaptation of survey instruments, data management, coding of students' responses, data analysis and reporting. They noted the benefits of strong peer-learning partnerships with representatives from PISA countries, and reported that their national assessments have been enhanced – particularly in the areas of student sampling, survey operations and data management – as a direct result of participating in the project. The countries would have appreciated more site visits by peer learning partners, the OECD and its contractors at key stages of implementation, but, unfortunately, the project had insufficient resources to provide this level of support.

6.6. Capacity building output achieved

Training in and support for assessment and analysis was successfully delivered to project country personnel in accordance with the PISA-D project design and in accordance with the schedule for the programme. As noted in Chapter 2, Strands A and B of the project were implemented successfully and in accordance with the plan and timeline agreed at the outset of activities and in accordance with PISA's technical standards. Also as noted in Chapter 2, one of the reasons for successful project implementation was the achievement of the capacity building output discussed in this chapter.

Key evidence for confirming that this output has been achieved is the successful implementation of PISA 2022 by Cambodia, Guatemala, Panama and Paraguay. These four PISA-D countries are implementing PISA 2022 with no additional support from OECD and its contractors or other development partners and are maintaining PISA's high technical standards and keeping up with the demanding PISA schedule with no problems.

Strand C of the project was designed to commence six months after the commencement of Strands A and B as it was expected that the managerial burdens for the countries and the contractors would be too great if they were implementing the in-school and out-of-school surveys at exactly the same time. Given the challenging nature of Strand C and the uniqueness of the approach – this was the first time that the National Centres had engaged in a household survey – the capacity building output was fundamental to the project's achievement.

7. PISA-D's contribution to the Sustainable Development Goals (SDG) - Agenda 2030

At the 70th Session of the United Nations General Assembly in September 2015, Member States adopted the 2030 Agenda for Sustainable Development. At the heart of the Agenda are 17 Sustainable Development Goals (SDGs), including SDG 4, which covers education.

Developed through an inclusive intergovernmental process, the 2030 Agenda integrates the social, environmental and economic pillars of sustainability with peace and security objectives. The Agenda in general, and SDG 4 in particular, bring together two earlier and parallel development strands:

- The environment agenda, known as Agenda 21, which was adopted at the 1992 UN Conference on Environment and Development, also known as the Rio or Earth Summit. Chapter 36 of Agenda 21 focused on the need for education for sustainable development.

- The poverty reduction agenda, known as the Millennium Development Goals (MDGs), adopted in 2000. MDG 2 focused on universal primary completion, while MDG 3 aimed at achieving gender parity in primary, secondary and tertiary education. In addition, the Education for All agenda also focused attention on goals related to early childhood, skills, adult literacy and education quality.

7.1. The Education Sustainable Development Goal (SDG 4)

The 17 Sustainable Development Goals (SDGs), adopted by the 70th General Assembly of the United Nations in 2015, are a universal call for action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity. The fourth SDG aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". SDG 4 is to be achieved by meeting ten targets, which together represent the most comprehensive and ambitious agenda for global education ever formulated.

Representatives of the global education community embraced SDG 4 by signing the Incheon Declaration at the World Education Forum in May 2015. The Declaration's 10 targets aim to support learning, in all its shapes and forms, which can influence people's choices to create more just, inclusive and sustainable societies. To advance progress towards SDG 4 and its targets, the global education community adopted the *Education 2030 Framework for Action* in Paris in November 2015.

The Framework for Action recognizes the UNESCO Institute for Statistics (UIS) as the 'official source of cross-nationally comparable data on education' and the Global Education Monitoring (GEM) Report as the 'mechanism for monitoring and reporting on SDG 4 … as part of the overall SDG follow-up and review'.

Following the Synthesis Report of the UN Secretary-General, a set of global indicators were proposed by the Inter-Agency and Expert Group on SDG Indicators and adopted by the UN General Assembly in July 2017. Data are published in the annual SDG Report. There are 11 global indicators for SDG 4, and the UIS is responsible for compiling data for 9 of them – the other custodian agencies are UNICEF (for 4.2.1) and the OECD (for 4.b.1).

As the 11 global indicators do not capture the full scope of the education agenda, a total set of 43 thematic indicators, including the 11 global indicators, constitute the SDG 4 monitoring framework. These are endorsed in the *Education 2030 Framework for Action*.

PISA-D and its contribution to the preparation of the SDGs

At the outset of PISA-D, the OECD established engagement with LMIC and partners for peer-to-peer analysis and learning opportunities to support the UN-led post-2015 process. This process began in 2013 at the same time as the development work commenced on PISA-D, as discussed in Chapter 1. Through its enhancement of PISA, the PISA-D initiative is designed to inform and support the monitoring, reporting and achievement of the education SDG and its related targets and indicators, particularly those related to learning outcomes.

PISA-D and the OECD's plans for mainstreaming the outputs of the project in future cycles of PISA is a key contribution to the continuation of these efforts, and an embodiment of international collaboration in support of the measurement and monitoring of learning outcomes in the context of the education SDG.

7.2. The OECD and the SDGs

Since the UN Summit in September 2015 that approved the **2030 Agenda for Sustainable Development**, the OECD has been working to help its Members and Partners adopt and adapt the **Sustainable Development Goals (SDGs)** and its associated targets. This has built on its long-standing work on well-being, and an OECD study Measuring Distance to the SDG Targets. The OECD is an observer on the UN Inter-agency and Expert Group on SDG indicators and is represented on the High-level Group for Partnership, Coordination and Capacity-Building. The OECD has been particularly active in supporting the development of an indicator framework to monitor the SDGs and fill in statistical gaps through methodological work.

The OECD has already provided data on agreed indicators for more than half the SDGs, either directly or in collaboration with other international agencies. For example, the OECD is working with UNDP through the Global Partnership for Effective Development Co-operation to contribute evidence and data to assess progress on SDG 5 on gender equality and women's empowerment, and SDG 17 on multi-stakeholder partnerships for the achievement of the goals (targets 17.15 on country's policy space and leadership and 17.16 on progress in multi-stakeholder development effectiveness monitoring frameworks). The OECD and the UNDP work together to feature this Global Partnership as evidenced through mandated follow-ups and review processes of the 2030 Agenda and Financing for Development, including joint contributions in related reports.

7.2.1. OECD and the Education SDG

The OECD has been working closely with UNESCO in support of the Education Sustainable Development Goal (SDG 4) and is a member of the UNESCO-led SDG Education 2030 Steering Committee with its associated bodies, including the Technical Co-operation Group (TCG), the Global Alliance to Monitor Learning (GAML) and the Global Education Monitoring (GEM) Report Advisory Board. The OECD's largescale survey data of the learning outcomes for 15-year-olds (PISA) and adults

(PIAAC) provides comparable and robust tools to measure progress towards SDG Targets 4.1 and 4.6[11].

The OECD has been a key partner of UNESCO and the other co-convening agencies in developing the education SDG framework, and works closely with UIS in the development of indicators that will be used to measure progress towards SDG achievement. In turn, UNESCO, UIS and the World Bank have partnered with the OECD in support of the PISA-D initiative. Since 2016, the OECD flagship publication, *Education at a Glance,* has reported data to help OECD and Partner countries assess their progress towards achieving the education SDG targets. The OECD has recently issued a co-signed letter with UNESCO and the UNESCO Institute of Statistics to all OECD Member and Partner countries confirming the collaboration between the three agencies in respect to SDG monitoring.

The OECD, UIS and the World Bank are working together and with other key practitioners, policy makers, researchers, representatives of governments, civil society organisations, funders, UN agencies, and other stakeholders committed to improving learning outcomes in all countries – particularly LMIC.

7.2.2. PISA, PISA-D and the SDGs

The SDG 4 agenda has rightly shifted the focus from the quantity (e.g. the number of children in school), which was a feature of the Millennium Development Goals that preceded the SDGs, to the quality and equity of education. Quality (i.e. achievement) and equity (i.e. fairness and inclusiveness) is harder to measure than simply the number of children in school; it requires reliable, relevant and useful learning data. Moreover, in order to serve the purpose of monitoring progress towards SDG 4, this learning data needs to be internationally comparable. During 2017 and 2018 considerable progress was made in the GAML, TCG and SDG Education 2030 Steering Committee to agree on a common minimum standard for reading literacy and maths literacy.

By including PISA and PISA-D data in the United Nations' global indicator framework (UNESCO Institute for Statistics, 2019[12]; United Nations Statistics Division, 2019[13]), the global community has recognised the role of PISA and PISA-D in monitoring progress towards the SDG for education over the next decade. PISA data are used in particular for monitoring progress in the proportion of children and young people who, at the end of lower secondary education, have achieved at least minimum proficiency in reading and mathematics (SDG global indicator 4.1.1c). PISA provides both a way of defining what "minimum proficiency level" means[14], through its described scale of proficiency, and a

---

[11] Target 4.1 By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes; Target 4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy

[12] **UNESCO Institute for Statistics** (2019), *Quick Guide to Education Indicators for SDG 4*, UNESCO Institute for Statistics, https://unesdoc.unesco.org/ark:/48223/pf0000265396.

[13] **United Nations Statistics Division** (2019), *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*, https://unstats.un.org/sdgs/indicators/indicators-list/

[14] Level 2 proficiency in PISA, the SDG 4 benchmark for reading and mathematics at the end of lower secondary education, is already used as a normative benchmark in many countries. For example, the European Union's strategic framework for co-operation in education and training

way of measuring this proportion, in an internationally comparable manner, amongst students who are close to the end of lower secondary education (or have recently completed lower secondary education). PISA-based indicators are also used to measure how close countries are to meeting other SDG targets, particularly those related to equity and education for sustainable development.

As Figure 7.1 shows, in B-S-J-Z (China), at least 90% of students attained Level 2 or above in reading in 2018, while in Dominican Republic, Kosovo, Morocco, and the Philippines, less than 30% of students attain this level of proficiency.

Figure 7.1. Proportion of 15-year-old students achieving at least minimum proficiency in reading (PISA Level 2 or above) in PISA 2018
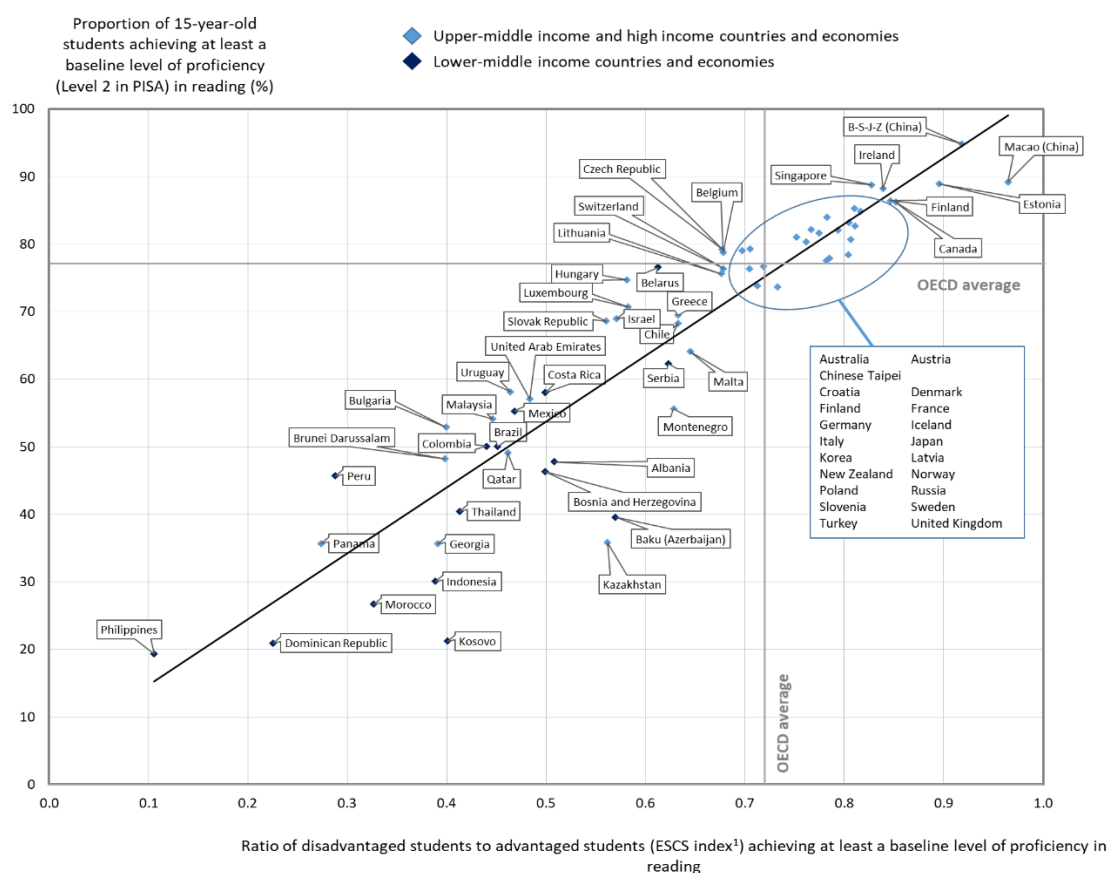
Note: 1. ESCS refers to the PISA index of economic, social and cultural status. Honk Kong (China), Netherlands, Portugal and United States data did not meet the PISA technical standards in 2018, and therefore are not displayed in the figure.
2. High-income countries and economies are defined here as those with a per capita GDP (PPP adjusted) above USD 20 000 in 2018.
Source: Add the source here. If you do not need a source, please delete this line.

(known as ET 2020), established in 2009, states: "By 2020, the share of low-achieving 15-year-olds in reading, mathematics and science should be less than 15%" (as measured by the proportion of 15-year-old students performing below Level 2 in PISA) (European Council, 2009)" http://dx.doi.org/10.3000/17252423.C_2009.119.eng.

As Figure 7.1 shows, lower-income countries participating in PISA are associated with higher proportions of students who do not attain the baseline level of proficiency in reading. About 20% of students in OECD countries (23% in 2018), on average, are low performers in reading. This proportion has remained stable since 2009. On average across OECD countries, the gender gap in reading in favour of girls narrowed by 12 points between 2009 and 2015; boys' performance improved, particularly among the highest-achieving boys, while girls' performance deteriorated, particularly among the lowest-achieving girls.

Figure 7.2 shows the ratios of the proportions of students achieving at least a baseline level of proficiency in reading[15] by gender, location (urban or rural) characteristics and socio-economic status (based on the PISA index of economic, social and cultural status [ESCS]). Among 15-year-old students, there are usually as many boys as girls who achieve at least proficiency Level 2 in mathematics, and more girls than boys who achieve Level 2 in reading. However, in the majority of OECD and Partner countries, their performance remains strongly determined by their school's location. Students who attend urban schools (located in communities with over 100 000 inhabitants) are more likely to outperform those who attend rural schools (located in communities with fewer than 3 000 inhabitants). Urban students tend to perform better because they go to schools that are usually larger and more likely to attract a larger proportion of qualified teachers. They are also more likely to come from a socio-economically advantaged background, which is directly linked to their performance in PISA (OECD, 2013[6]).

---

[15] The proportion of the likely most disadvantaged group is in the numerator of the ratios. A ratio equal to 1 indicates parity between the two considered groups. A value less than 1 indicates a disparity in favour of the likely most advantaged group and a value greater than 1 a disparity in favour of the most disadvantaged group.

Figure 7.2. Inequalities in achievement by gender, wealth and location (2018)

Proportion of 15-year-olds achieving at least PISA proficiency Level 2 in reading

Distance to 1 of the ratio of the proportions
of students achieving at least a baseline level
of proficiency in reading by gender, wealth
and school location[1]
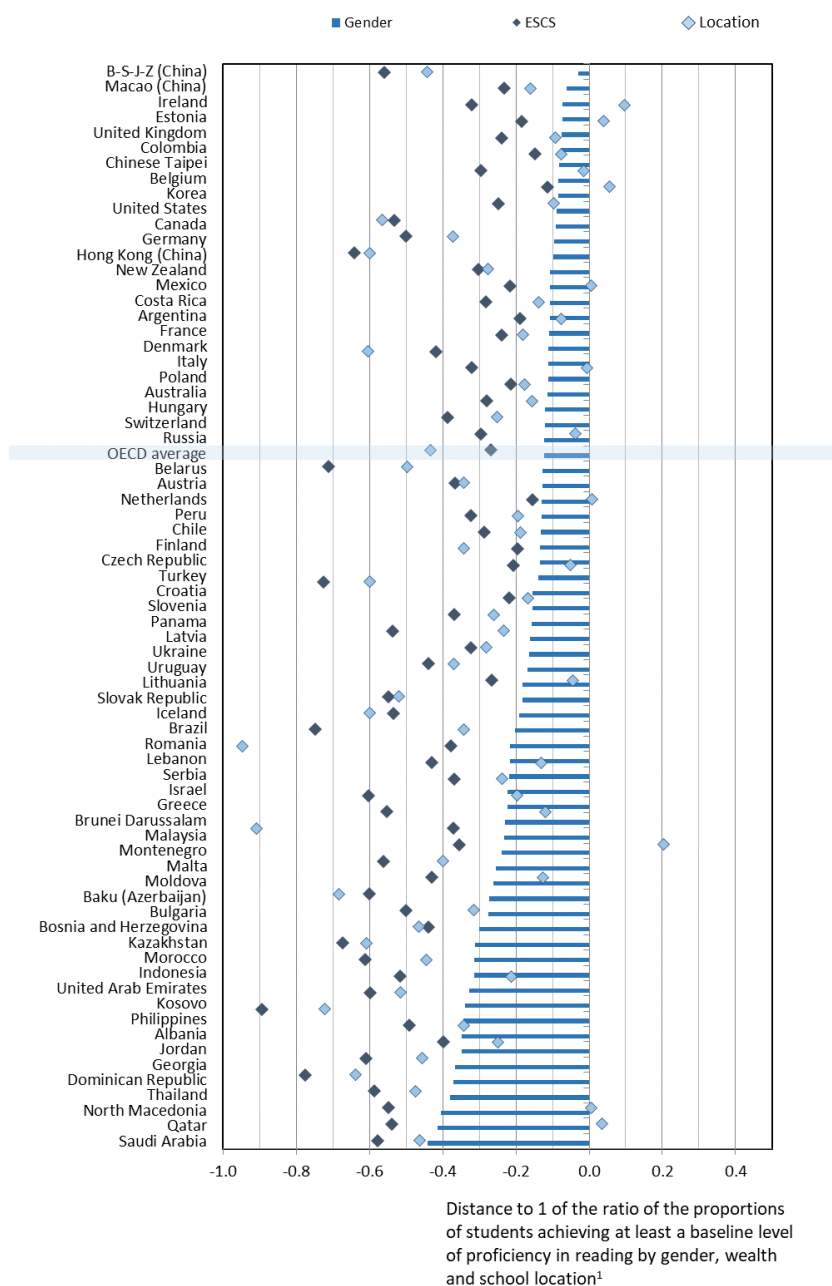
Note: Countries and economies are ranked in ascending order of the distance to 1 of the ratio of the proportions of students achieving at least a baseline level of proficiency in reading in reading, by gender.
1. The proportion of the likely most disadvantaged group is in the numerator of the ratios. A ratio equal to 1 indicates parity between the two considered groups. A value less than 1 indicates a disparity in favour of the likely most advantaged group and a value greater than 1 a disparity in favour of the most disadvantaged group.
Source: OECD, PISA 2018 Database.

The performance gap between students from different socio-economic backgrounds remains a reality in all countries, both in reading and mathematics. Even in those countries where equality is (almost) met along the three dimensions, the proportion of young people achieving PISA Level 2 in reading is between 10% and 20% smaller among the most disadvantaged students. Even more worrying, levels of socio-economic inequality have not changed since 2006 in the majority of countries. Figure 7.3 shows that in a few countries, including Finland, Iceland, Thailand, and Korea, the disparity between students in the top and bottom quartiles of the PISA index of socio-economic status grew even larger between 2006 and 2018. However, PISA results show that inequity of opportunity is not set in stone, and that selected school systems succeeded in becoming more equitable over a relatively short period (OECD, 2017[7]). This is the case in Austria, Denmark, Germany, Lithuania, Montenegro, Norway, Russia, Serbia, Slovak Republic, and other countries, where the performance gap between the quartiles of socio-economic status narrowed significantly between 2006 and 2015. However, large differences in performance between disadvantaged and advantaged students remain in these countries.

Figure 7.3. Trends in socio-economic inequalities in achievement (2006, 2018)

Proportion of 15-year-olds achieving at least PISA proficiency Level 2 in reading

Ratio of the proportions between bottom quarter and top quarter of the PISA index of socio-economic status
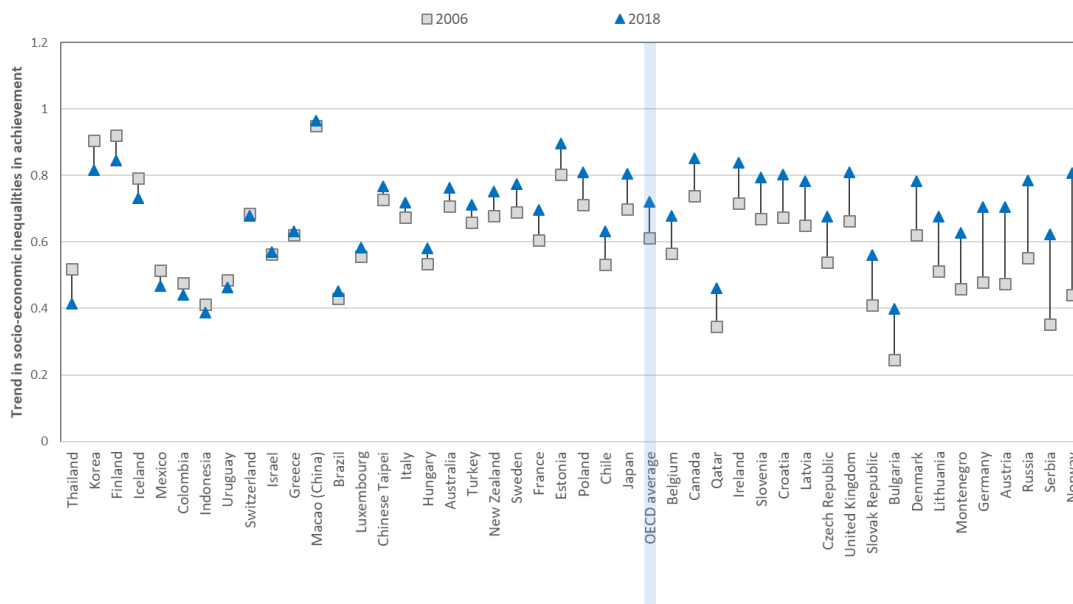
Note: Countries and economies are ranked in ascending order of improvements in equity since 2006.
Source: OECD, PISA 2018 and PISA 2006 Databases.

PISA shows that in many countries, no matter how well the education system performs as a whole, socio-economic status continues to predict students' performance. However, PISA also consistently shows that high performance and greater equity are not mutually exclusive. Being able to improve the performance of all students, regardless of their background, is necessary for countries to become high performers and attain the SDG 4 targets.

Making PISA more accessible and relevant for a wider range of countries

PISA results highlight differences in education quality between high-income and middle-income countries: students in middle-income countries perform well below the OECD average (Figure 7.4) and their performance is concentrated at the lower levels of the PISA proficiency scales. The limited differentiation of performance at these lower levels constrains the knowledge and understanding of what these students can and cannot do. It also limits the analyses that can be done linking lower levels of learning with education policies and student's characteristics

Figure 7.4. Percentage of students scoring at Level 1 and below in reading in 18 middle-income countries, PISA 2018
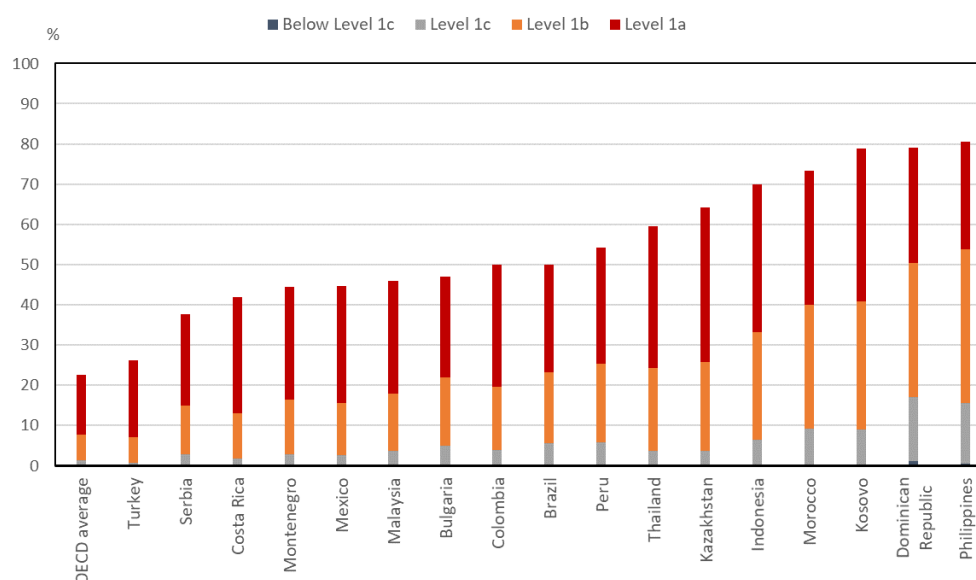
Note: Countries and economies are ranked in ascending order of proportion of students achieving Level 1 or below in reading, in 2018.
Source: OECD, PISA 2018 Database.

Some of the contextual factors measured by PISA are unrelated to differences in performance in middle-income countries; and PISA does not adequately reflect some of the contexts unique to these countries (Lockheed, Prokic-Bruer and Shadrova, 2015). Because many 15-year-olds in middle-income countries do not attend school, coverage can be as low as 50% (Spaull, 2017). In addition, some middle-income countries have encountered financial, technical, and institutional difficulties in implementing the assessment and using PISA data.

PISA for Development (PISA-D) has made PISA more accessible and relevant to a wider range of countries. It has extended the PISA test instruments to measure a broader spectrum of performance, particularly at Level 2 and below. This is facilitating greater knowledge and understanding of what lower performing students can do. It has developed contextual questionnaires and data-collection instruments to capture the diverse situations in LMIC. In addition, PISA-D has established methods and approaches to include out-of-school

youth in the assessments – thus potentially offering a continuum between PISA and the OECD's Survey of Adult Skills (PIAAC) in terms of target populations and contributions to global indicators of SDGs, and it has built capacity in the participating countries to manage and use the results of large-scale student assessments.

While the PISA-D test design and items target the lower levels of performance, the assessment is linked to the whole of the PISA framework for comparability. This link is established using data from some of the PISA 2015 trend questions. PISA-D provides a way of measuring differences in performance at the low end of the proficiency scale for each subject tested (reading, mathematics, and science) even as it measures performance at the higher levels. The PISA-D cognitive test lasts two hours, as does the main PISA test, and the assessment is conducted in accordance with PISA's technical standards.

There is already evidence showing that the PISA-D project has helped build the capacity of the participating countries to manage and make good use of large-scale assessments. As noted above, four PISA-D countries (Cambodia, Guatemala, Panama and Paraguay) are successfully implementing PISA 2022 with no additional support from the OECD or its contractors or other development partners. With the enhanced instruments and approaches from PISA-D made available in the main PISA test, it has been possible to incorporate a large group of new participants in the 2021 cycle of PISA. The project is on track to provide important insights into quality and equity in education in the participating countries, and to allow more countries to participate in PISA – all of which will help measure global progress towards attaining SDG 4 without excluding out-of-school youth.

PISA shows that in many countries, no matter how well the education system performs as a whole, socio-economic status continues to predict students' performance. However, PISA also consistently shows that high performance and greater equity are not mutually exclusive, as shown in the figure. Being able to improve the performance of all students, regardless of their background, is necessary for countries to become high-performers and to achieve the SDG 4 targets.

7.3. Expanding the coverage of PISA to better monitor progress towards achieving SDG 4

PISA-D is making PISA more accessible and relevant to a wider range of countries. It is extending the PISA test instruments to measure a broader spectrum of performance, particularly at Level 2 and below. It is also developing contextual questionnaires and data-collection instruments to capture demographic information about participating LMIC. The aim is to allow more countries to participate in PISA – and thus to be able to measure more widely global progress towards attaining SDG 4.

7.3.1. The mutual benefits of wider participation in PISA

While facilitating better monitoring of progress towards SDG 4 was a major objective of PISA-D, bringing more countries into the Programme also gives OECD Members access to more diverse perspectives on education reform and a wider range of policy experience and solutions. The participation of Partners in the work of the PGB has thus benefitted the Programme, as it has been and continues to be an opportunity for including more diversity in policies and practices, enriching analyses by having a greater range of points of comparisons and also increasing the opportunities for peer learning. For example:

- having high performing Partners has enriched the analyses on high performance beyond Members only

- having rapid improvers, such as Peru, has allowed for a description of the changes in policies and practices that may help explain these improvements

- having LMIC Partners, such as Latvia (an OECD member since 2016) and Croatia, allowed to develop a narrative about different relationships between resources and performance according to Partners' income level and to better identify the minimum levels of resources that are necessary to make an impact on performance.

Engagement in PISA from Partners also provides OECD Members with a means to positively influence education performance in Partners whose economies and labour forces are increasingly integrated with their own. It should also be noted that several Partners are actually receiving aid for education from OECD Members and thus participation in PISA facilitates effective monitoring of education outcomes for both the providers of development co-operation and the recipients.

Partners can benefit from access to PISA to benchmark themselves against global standards and obtain high-quality data on education and the distribution of learning outcomes throughout the population. They also benefit from the OECD's policy know-how, and from the peer-learning and capacity-development opportunities provided by participation in the assessment and contact with the OECD's networks of experts. Particularly for LMIC Partners, PISA participation helps to develop and enhance the capacities needed to conduct national and international assessments and PISA data may be the most reliable way of collecting national level data in the absence of robust national assessments. Engagement in PISA also enables Partners to generate positive domestic pressures for catch-up.

The important mutual benefits of Partner engagement in the work of the OECD and in PISA are based on the three imperatives outlined below.

The education imperative

Recent decades have witnessed a significant shift in the global balance of education skills that is only likely to intensify in the coming years. Demographic trends combined with progress in expanding access to education mean that non-Members account for an increasing share of the global talent pool; already, tertiary graduates in Key Partners outnumber those in OECD Members combined. While learning outcomes remain on average higher among OECD Members, PISA shows that some of the world's top performing education systems – and most rapid educational improvers – are outside the OECD. The pace and pressure for change in the emerging world has generated new policy approaches and innovative solutions from which all countries and economies can learn. If the OECD is to retain its influence as a leading source of education good practice, it will need to draw on more diverse experiences and strengthen its engagement with non-Members that have successful policies to share.

The provision and demand for education are also becoming more global, and education systems have been experiencing common challenges and opportunities. Globalisation, technology and migration have made national education systems more porous. The increase in migration and the refugee situation of recent years has brought into the policy spotlight the increasing diversity of OECD classrooms and the impact that education systems in non-Members can have on schooling within the OECD.

Together, these trends point to the importance of being able to monitor the progress of learning and skills development globally, of having global networks where countries and economies can agree common standards and share good practice, and of monitoring and anticipating changes in the global education landscape.

The economic imperative

The increasing interconnectedness of economies and the broader range of countries driving global growth are decisive factors influencing the OECD's approach to global relations. Notably, the shift in the world's centre of economic gravity has led to determined efforts to strengthen engagement with major emerging economies, in particular the Key Partners (Brazil, China, India, Indonesia and South Africa), and to create more inclusive models of global economic governance. These efforts have achieved important results, with the establishment of joint work programmes between the OECD and Brazil, China and Indonesia.

At the same time, the OECD has positioned itself as the technical partner of choice for the G20 forum of major developed and emerging economies, providing advice on a wide range of issues, from tax reform to employment and social policy. All but two G20 countries (India and South Africa) today participate in PISA, and Brazil is an Associate of the Programme. However, data on several important non-Members remains limited and more needs to be done to integrate non-Members into OECD policy dialogue on skills. By deepening its knowledge base on major emerging economies the OECD would be in a better position to influence this discussion, as it has done so effectively in other areas.

It is important that education and skills feature prominently in these partnerships if they are to achieve their goal of strengthening the foundations for stable global growth. Recent OECD economic surveys of Key Partners have identified gaps in education participation and outcomes as one of the most significant structural constraints to sustained growth in the emerging world. From productivity and formal employment to economic diversification and reduced dependency on commodities: progress on a range of economic indicators will be held back unless emerging economies can develop higher quality skills and widen their talent base.

The achievement of global sustainable development is related with propelling national education reform and strengthening the structural conditions for international growth. Strong education systems in non-Members can accelerate progress towards higher income status and reinforce supply chains and global markets, and more generally, the benefits to global peace, security and prosperity of having all countries with improved education systems are significant. In the long term, there are greater benefits of this kind from improving the education systems of LMIC non-Members, than from raising the performance of OECD Members.

The development imperative

As noted already in this chapter, the adoption of the *2030 Agenda for Sustainable Development* and of the Sustainable Development Goals (SDGs) by the United Nations (UN) General Assembly in September 2015 has made education a worldwide priority with SDG 4 (the Education SDG): "*Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*" All UN member states committed to this universal agenda that calls for countries at every level of development to work together towards common goals. For the OECD, this means an acceleration of the transition envisaged in the 2012 Strategy on Development towards more inclusive policy dialogue and the participation of a wider array of countries in its activities, as outlined in the OECD's Action Plan on the SDGs [C(2016)68]. Nowadays, OECD support to global sustainable development goes well beyond the OECD development cluster and is increasingly integrated in the work of substantive programmes, including education and skills.

The SDGs are strongly focused on learning outcomes, where the OECD is a leading source of comparative data and policy know-how and on the inputs and processes needed to ensure quality education. Indeed, PISA has been included in the SDGs global indicators framework that will be used to monitor progress towards the Education SDG by 2030 as the principal source of data for measuring progress against SDG global indicator 4.1.1 (c): Proportion of children and young people: at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.

PISA is the world's most inclusive metric on student outcomes and one of the most global instruments of the OECD already, with 78 OECD Members and non-Members involved in the 2018 round. PISA is particularly valuable not just for monitoring SDG progress but also for informing policy and funding decisions that will help countries achieve the SDGs.

The centrality of education to sustainable global development calls for more proactive efforts by the OECD to understand the distinct issues facing Partners, offer support to national reform efforts, and contribute to the global architecture and strategies that the OECD has helped to put in place for the monitoring and achievement of the Education SDGs. The PISA Development Strategy approved by the PGB in June 2016 highlights the importance of accommodating an increasing number of LMIC non-Members in PISA while maintaining the technical integrity and standards of PISA. Through this strategy and the PISA-D initiative PISA is being enhanced to enable more non-Members, especially LMIC ones, to participate. By facilitating peer-to-peer learning among new and established participants, PISA-D also demonstrates how the OECD can be a catalyst of stronger bilateral and multilateral cooperation among countries and economies, strengthening development partnerships.

With the adoption of the *2030 Agenda*, it is clear that there will be both increased opportunity and demand for the OECD to engage in global education development. While the OECD is an observer in most United Nations processes, the Organisation is a full member[16] of the UNESCO-led SDG-Education 2030 Steering Committee, which is the main global consultation and coordination mechanism for education in the *2030 Agenda for Sustainable Development*. It is, thus, important for PISA to contribute to this increasing engagement in the sustainable development agenda and to determine how best to apply an SDG lens to its programme of work.

7.4. Output achieved

The project was successful in establishing engagement with LMIC and partners for peer-to-peer analysis and learning opportunities to support the UN-led post-2015 process. PISA and results from the PISA-D pilot were considered in the UN-led discussions of the post-2015 process and are being used to monitor progress towards achieving SDG 4. In particular, the data from PISA and PISA-D are being used by the UN system to measure progress against global indicator 4.1.1.c for more than 90 countries – this is one of the highest levels of coverage for the learning outcome indicators included in the SDG4 monitoring framework.

---

[16]The Council has endorsed the OECD's role in SDG4, including its membership in the Committee. Membership includes participating in the meetings of the Committee, leading the Committee's Working group on Policy and strategy, participating in the Committee's Working group on monitoring and reporting, participating in the Committee's Technical Coordination Group, the Committee's Global Alliance to Monitor and other related technical and working groups.

8. Conclusions, mainstreaming, lessons learned, outcome, impact and next steps

The instruments developed through the PISA-D initiative provide a finer-grained view of low-performing students and better measure factors more strongly related to student performance in LMIC. In addition, the methods for assessing out-of-school youth piloted in PISA-D show considerable promise as a means for countries without universal access to secondary education to still obtain a complete picture of the skills of all their 15-year-olds. PISA-D also established a model to build capacity in the participating countries for managing large-scale student learning assessments and using the results to support national policy dialogue and evidence-based decision-making.

One of the most significant achievements of PISA-D was the active engagement of the participating countries in the analysis and use of the data. A lack of this engagement by LMIC in PISA had been highlighted as an issue to be addressed during the preparation of PISA-D.

A further significant achievement of the project is its contribution to the development of PISA itself through the enhancements it has made to the instruments, especially the greater granularity of the cognitive instruments at lower levels of performance and the increased relevance of the contextual data collection instruments for a wider range of countries.

At the beginning of the initiative, countries indicated that they were participating in PISA-D in order to:

- gain an understanding of how their students' performance compares internationally and, in particular, how it compares with student performance in countries facing similar challenges

- develop insights into how to help students learn better, teachers to teach better and school systems to operate more effectively

- build capacity to conduct large-scale learning assessments, and analyse and use the results to support national policies and evidence-based decision making.

Seven of the countries participating in the PISA-D school-based assessment – Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Senegal and Zambia – reported results in December 2018 in individual national reports; Bhutan reported its results in March 2019. As the initiative completes its final phase, it is a good time to reflect on whether the countries' expectations have been met.
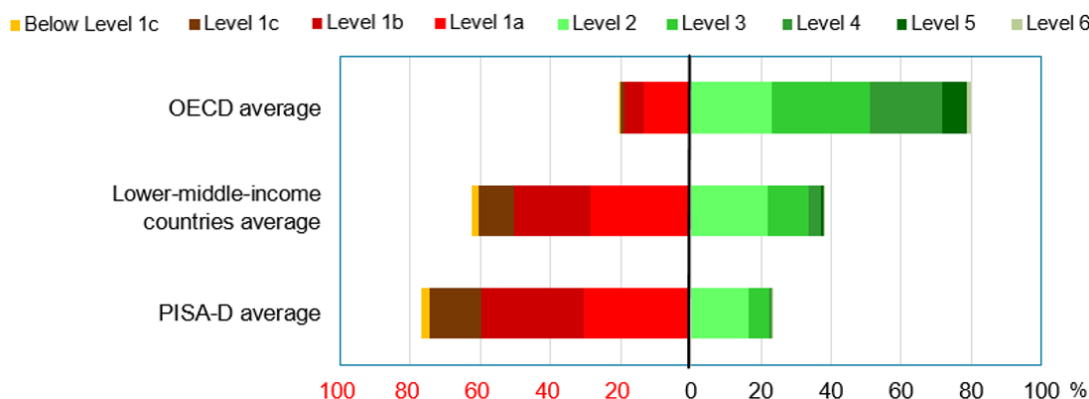
8.1. Conclusions – expectations have been met

PISA-D was successful in making the assessment instruments more relevant to LMIC while still being able to report results on the main PISA scale, thus facilitating international comparisons on all of the variables covered by PISA (see Figure 8.1 below). These variables and factors include student performance, educational attainment, health

and well-being, attitudes towards school and learning, the learning environment, learning time, the quality of instruction, family and community support, and resources devoted to education.

Figure 8.1. Students' proficiency in reading, PISA 2015 and PISA-D

8.1.1. Policy insights

PISA-D assessment results provide countries with a solid database that can help them refine policy priorities and set new goals or targets to improve their education systems. The data collected have a lot to say about the allocation of resources and its implications for equity. With reliable data on gaps in access and differences in outcomes between groups of children and young people, countries can determine whether poor and marginalised populations are given equal opportunities to succeed at school and beyond. The challenge for countries is to maintain a focus on these goals or targets, and to track progress towards them by participating in future cycles of PISA and other relevant studies. With four PISA-D countries already participating in PISA 2022 and three others confirming their intention to participate in PISA 2025, this challenge appears to be manageable in the medium term at least.

8.1.2. Capacity development

PISA-D has helped the participating countries build their capacity to manage large-scale assessments and make use of the results in support of national policy dialogue and education policy making. When asked about the usefulness of the capacity building provided through PISA-D, countries largely praised the continuous support they received through every phase of the project. Countries also highlighted the pertinence of the training provided, including sampling, translation/adaptation of survey instruments, data management, coding of students' responses, data analysis and reporting. They noted the benefits of strong peer-learning partnerships with representatives from PISA countries, and reported that their national assessments have been enhanced – particularly in the areas of student sampling, survey operations and data management – as a direct result of participating in the project. The PISA-D countries appreciated the needs analyses conducted at the beginning of the project as both a useful assessment of the national team's strengths and weaknesses and as a benchmark for assessing how far they have come in capacity building since the project began.

8.1.3. PISA for Development construct validity

The OECD has enhanced its PISA instruments through the PISA-D initiative to provide a finer-grained view of low-performing students and to better measure factors more strongly related to student performance in LMIC as described in the PISA-D Assessment and Analytical Framework.

The evidence from the implementation of the enhanced PISA-D instruments against the theory that underpins them as set out in the PISA-D Framework, and the PISA technical standards that govern the assessment (as documented in the PISA-D Technical Report) all confirm the construct validity of the PISA-D test scores for the proposed uses of the tests. The cognitive test results show that it really was worth the effort of developing level 1.c.

The educational prosperity approach

To make PISA more relevant for LMIC, PISA-D has adopted an educational prosperity approach. This approach is the basis of the PISA-D Framework and considers the conditions for the success of education systems throughout a person's lifetime. It identifies a set of four key outcomes, called "Prosperity Outcomes", for each stage of schooling and child development: educational attainment; academic performance; health and well-being; and attitudes towards school and learning. The prosperity approach also identifies a set of family, institutional and community factors, called "Foundations for Success", that influence these outcomes: inclusive environments, learning time, quality instruction, family and community support, and resources (see Figure 8.2 below).

Figure 8.2. The educational prosperity model
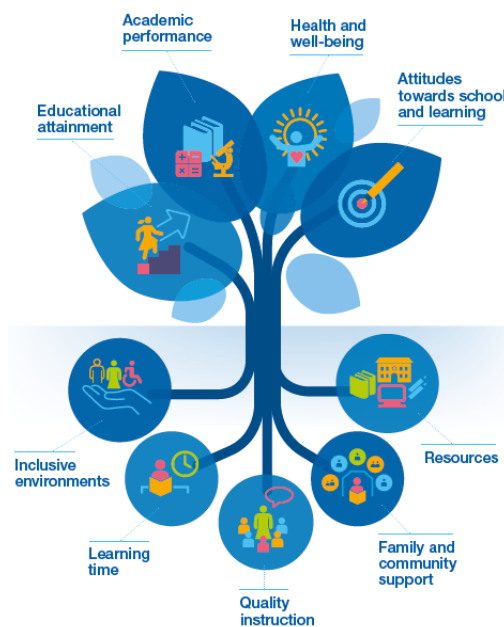
The approach has three explicit links to national and local policy and practice.

- First, it allows countries to set goals or targets for improvements in the Foundations for Success, based on PISA-D data, at all levels of the system, from the education minister and his or her staff to front-line educators, students and parents. The challenge for countries over time is to maintain a focus on these goals or targets and to track progress towards them by participating in future cycles of PISA and other relevant studies that may generate comparable results to PISA.

- Second, the data collected has immediate implications for education policies concerning the allocation of resources and its implications for equity. With reliable data on differences, between groups of young people, in outcomes and access to the Foundations for Success, countries will be able to determine whether poor and marginalised populations are given equal opportunities to succeed at school and beyond.

- Third, the data collected will enable countries to set goals and targets for improvements in their Foundations for Success that are consistent with the framework of the SDG for education and monitor progress towards them. PISA-D provides an infrastructure for analysing relationships between trends in outcomes and policy changes. The descriptive evidence from PISA complements policy evaluations and more qualitative assessments of the implementation of policy reforms.

Validating the assessment

The field trials of the PISA-D cognitive tests in 2016 provided information about the data collected and survey operations, assessed the quality of the test items, and helped determine the reliability and comparability of the PISA-D and PISA scales. Using the background questionnaires during the trials allowed for selecting items for the final instruments based on their psychometric properties. PISA's technical standards were applied at every stage of the project. The main survey data collection is subject to a strict adjudication process, particularly for the sampling and translation/adaptation parts of the implementation. The way the sample is selected and the instruments are localised in each country is an integral part of the evidence supporting the construct validity of the intended uses of the assessment. (Invalid uses of PISA data include ranking schools within a country and measuring the value-added of teachers.)

During the analysis phase of the project, analysts took an in-depth look at the results, including the functionality of all test items, and confirmed that the instruments measure what they purport to measure against the PISA-D Framework. Indeed, the analysis validated the assumptions that shaped the PISA-D Framework. As a result, participating countries can be confident that the assessment results will provide relevant data that can inform decisions concerning national policies.

8.2. Mainstreaming and lessons learned

Mainstreaming, or broad adoption, of the outputs of the PISA-D project in the main PISA programme is critical to the initiative's success. While scaling-up and mainstreaming of these outputs is currently a work in progress, it is possible to report on achievements so far and to also summarise some of the key lessons learned from the pilot in each of the five output areas.

8.2.1. PISA's 2021 instruments will be better suited for assessing lower levels of performance and socio-economic status.

The PISA 2022 cycle will focus on mathematics as the main cognitive domain. The results of the PISA-D initiative are being integrated into the mathematics framework and item development for this cycle for the computer based assessment (CBA), including lower performance levels and items that measure these levels of proficiency. For countries conducting the computer-based assessment, the cognitive instruments are moving towards adaptive testing procedures in order to improve the quality and accuracy of the data collected. This will allow for far more differentiation of student performance at different levels on the proficiency scales, especially at the bottom and top of those scales.

Despite PISA's shift to a predominately computer-based assessment, paper-based forms will continue to be offered in 2021 to allow countries without adequate computer coverage to participate in the programme. The paper-based version will be the same as that used in PISA-D; at least 60% of the items will be at proficiency Level 2 and below (PISA describes six proficiency levels, with Level 6 the highest and Level 1 and below the lowest).

From 2021 onwards, the PISA questionnaires will incorporate some of the PISA-D questions, particularly those that extend the measure of student economic, social and cultural status, to better capture lower levels of parents' education and income, and risk factors of poverty that are more frequently found in LMIC. Adaptive-style questions are being introduced so that socio-economic status can be more accurately measured without significantly extending the time required to complete the questionnaire. Depending on the results of PISA-D, other questions on topics such as pre-primary attendance, language of instruction versus language spoken at home, quality of instruction, and how long students spend receiving different types of instruction during the school day might also be included.

8.2.2. Capacity-building support is being offered as an optional component of PISA

The PISA-D model of preparing and supporting new countries to participate successfully in PISA is being incorporated in PISA from the 2022 cycle onwards. Four new countries (El Salvador, India, Mongolia and Uzbekistan) have opted for capacity-building support from the OECD and its contractors in PISA 2022. This support is provided first through the three-stage process piloted successfully in PISA-D. The process begins with an analysis of a country's capacity to implement PISA and make use of PISA data, and includes planning to strengthen that capacity. In addition, countries opting for this capacity-building assistance are supported by the OECD and its contractors at each stage of the PISA cycle:

Six new countries (El Salvador, India, Jamaica, Mongolia, Romania and Uzbekistan) have also opted for training and assistance in data analysis, the interpretation of PISA results, report writing, and communication in PISA 2022 through a process that has been successfully piloted in PISA-D and repeated for seven countries in PISA 2018:

| Preparing stakeholders for PISA results | → | Analysis of PISA results and country report | → | Dissemination of results and evidence-based policy dialogue | → | Better policies for improved learning outcomes |

These types of support are designed to help LMIC, in particular, to overcome two potential barriers to their participation in PISA: a lack of capacity to implement the assessment, and a lack of experience in using PISA data and results.

8.2.3. An out-of-school assessment will be offered as an optional component.

PISA will also offer an out-of-school assessment as an optional module in future cycles. Piloted in PISA-D, this module measures the competencies of out-of-school 15-year-olds, providing a context for interpreting the in-school results for PISA-participating countries that have sizeable proportions of 15-year-olds who do not attend school. With this enhancement, PISA will be able to provide countries with important information about the human capital in the population as a whole, not just among those who have attained grade 7 or higher by the time they are 15 years old.

The OECD will also offer as an international option as part of a future PISA cycle an out-of-school assessment that is *linked to a household survey* such as UNICEF's Multi Indicator Cluster Survey (MICS) or USAID's Demographic Household Survey (DHS) thus reducing considerably the cost of identifying and locating eligible youth for the assessment.

The OECD is also working with key partners such as the World Bank, UNESCO and UNICEF to offer a shortened version of the PISA-D cognitive test as a learning assessment module that can be *integrated with a multi-topic household survey.* This test is designed solely to discriminate whether a respondent is above or below 406 points on reading and 419 points on mathematics – the SDG 4 benchmarks for minimum levels of proficiency. This integrated PISA module may be taken by a country as part of a future PISA cycle or a completely separate study. This PISA module would be targeted particularly at low-income countries with the aim of closing learning assessment data gaps for the monitoring of progress towards SDG 4.

Incorporating these enhancements from the PISA-D out-of-school assessment pilot into PISA will make the assessment more relevant to a wider range of countries, especially LMIC. This will also help monitor progress towards the Education SDG, which emphasises ensuring that all children and young people achieve at least minimum levels of proficiency in reading and mathematics.

8.2.4. Lessons learned

The following section summarises some of the key lessons learned from the pilot in respect of project implementation and in each of the five output areas.

Project implementation lessons

- PISA-D succeeded because it was a truly collaborative effort, bringing together experts from the participating countries, and steered jointly by their governments on the basis of shared, policy-driven interests.

- PISA-D also succeeded because experts from participating countries serving on working groups were able to link policy objectives with the best internationally

available technical expertise. By participating in these expert groups, countries ensured that the instruments were internationally valid and took into account the cultural and educational contexts in the participating countries, that the assessment materials had strong measurement properties, and that the instruments emphasised authenticity and educational validity.

- The NPMs in each country played a vital role in ensuring that the implementation of the survey was of high quality, and verified and evaluated the survey results, analyses, reports and publications.

Enhancement of the cognitive instruments

- PISA-D instruments work: they capture a wider range of student performance while ensuring that results are comparable to those of the main PISA test.

- The overall performance of 15-year-old students in all of the countries that participated in PISA-D, was much lower than that of students in OECD countries, and varies widely.

- PISA-D cognitive instruments have a greater number of test items at the baseline level of proficiency and below to give a fine-grained indication of the knowledge and skills of lower-performing 15-year-olds – valuable information for designing quality-improving programmes.

- A series of enhancements are being incorporated into the PISA instruments to better describe the performance and contexts of a wider range of students. These enhancements have enriched the PISA instruments to provide relevant policy information not only for LMIC Partners, but also for Members, who all have some proportion of their students who perform poorly or come from disadvantaged backgrounds and are an important focus for policy.

- Starting with Reading in 2018 with the inclusion of Reading Components, it is expected that proficiency levels 1.b and 1.c. will be gradually incorporated and that gradually all domains will include more items at the lower levels of performance (level 2 and below) to better measure the skills of the lowest performing students. In the case of the computer based assessment (CBA) this will be facilitated by adaptive testing.

- In PISA 2022 CBA the increase in lower level items will be concentrated in Mathematics, the major domain, and in the 2024 CBA the emphasis will be placed on Science. This new feature will benefit all participating Members and non-Members, as even high-performing countries and economies have some low-performing students.

- The PISA D paper based test, that has a higher concentration of lower level items in the three domains, will be offered as an option for new Partners in PISA 2022.

Enhancement of the contextual questionnaires

- PISA-D instruments work: they capture the diverse contexts found in LMIC while ensuring that results are comparable to those of the main PISA test.

- It is expected that the contextual questionnaires in PISA 2022 will seek to extend PISA's measure of economic, social and cultural status (ESCS) to better describe the contexts of lower income level students in all Partners and Members, and may

also incorporate other variables to improve understanding of factors associated with performance in low-and-middle-income contexts. To implement this and to be able to better respond to the different demands and interests of Members and Partners, the possibility of modularising groups of questions is being explored.

Incorporating the out-of-school into future PISA assessments

- The PISA-D out-of-school assessment shows how it is possible to reach a nationally representative sample of 14-16 year-olds who are out of school or in grade 6 or lower and to test these in their homes while ensuring that results are comparable to those of the main PISA test.

- The PISA-D out-of-school assessment also shows that a nationally representative sample is not the best approach in all countries and a range of sampling options are available to suit all purposes and budgets.

- The delivery of an out-of-school assessment via tablet computer in the homes of eligible young people is feasible, even in low-income countries. It includes a background questionnaire, and reading and mathematics tests for the young person, a background questionnaire for his or her parent/guardian, and observation questions for the interviewer conducting the survey.

- A National Survey Operations Manager is necessary for each participating country – this person has skills and experiences in survey operations that are different to those usually found in an NPM. He or she is responsible for all survey operations in the country, including preparing training materials, recruiting and training interviewers, and managing field operations, with the support of the OECD's international contractors.

- It is extremely challenging for a country to manage an in-school assessment at the same time as an out-of-school assessment – therefore, the two assessments should be staggered in future cycles of PISA.

- An optional module to incorporate out of school 15 year olds in the assessment may be offered in future PISA cycles, so participating Members and Partners who choose this option can obtain information about the skills acquired by all children, not just those who attend school. This option would be also suitable for Members that have high percentages of out of school 15 year olds, not only for Partners.

- The costs of identifying and locating target respondents for the out-of-school survey are prohibitively high. These costs can be reduced considerably by *linking* future PISA out-of-school surveys to household surveys. With this linking the identification and location of respondents is a by-product of the household survey and the National Centre's in-country costs and sampling tasks are thus greatly reduced. Such linking will also yield considerable benefit by providing a richer set of contextual data for each respondent

- The PISA-D out-of-school instruments, suitably modified, have the potential to be offered to countries as a PISA assessment module for delivering as part of a household survey; i.e. a shorter assessment that would still provide results that are comparable to those of the main PISA test. This PISA module can be linked to or integrated with a household survey, thus reducing considerably the costs to a country of obtaining a PISA score, albeit a limited score (proportions of

respondents above and below Level 2 in reading and mathematics) compared to full PISA participation.

Capacity building lessons

- Preparing countries for their participation in PISA-D through the three-stage process was effective.

- The technical training delivered to PISA-D countries has helped them complete the complex tasks required of them in order to advance from one phase of the project to the next while adhering to PISA's technical standards.

- The OECD and PISA-D contractors closely monitored countries' progress against clearly defined tasks at each stage of survey implementation, and this has proven essential for identifying countries in need of extra support and providing assistance in a timely manner.

- While the results of the field trial were positive and its goals were accomplished, each country faced its own challenges in conducting the school-based assessment, such as completing all field-trial tasks on time. The field trial helps countries better plan for the Main Survey and anticipate challenges.

- Providing countries with tailored support to reinforce their capacity and create the conditions in each country that are needed to successfully implement the assessment is essential to the success of the project.

- Learning-by-doing works: simply by participating in PISA-D with its well-established and high-quality procedures and technical standards, the countries have acquired valuable knowledge and understanding of how to manage a large-scale assessment.

- Peer-to-peer learning is an effective way to build capacity, especially when it takes place in the context of an on-going cycle of activity like the PISA assessment.

- To maximise their participation in PISA, support for analysis and reporting was offered to Partners as an option in PISA 2018 (this took place during 2019 and was very successful) and PISA 2022 (six countries have already signed up for this option). This option will be offered to new Partners and to Partners that have already participated in previous cycles but would like more support in producing national reports. In addition to this, PISA is constantly reflecting on how to improve the analysis and communication of international results and in the context of the increasing numbers of participants – how to best present country comparisons in the international reports – is one of the elements being taken into account.

- While it is likely that in the near future most LMIC Partners will be able to manage computer based assessment, in the immediate term some of them may face challenges, especially newcomers. For this reason the PISA 2022 cycle has offered the "PISA D paper based assessment", (which has a higher concentration of items at level 2 and below).

8.3. Outcome and impact of PISA-D

The development hypothesis that informed the design of PISA-D was that if the five outputs of the project discussed in this report were achieved, then it would result in the achievement of the following outcome:

- Increased numbers of LMIC use PISA assessments from 2021 onwards to monitor progress towards national improvement targets, to comparatively analyse factors associated with student outcomes, for institutional capacity-building and for tracking international education targets within a post-2015 framework.

The logic of the hypothesis continued with the assertion that if this outcome was achieved, then the project would have the following impact:

- The majority of countries – high income and LMIC – are able to effectively monitor progress towards national and international goals and targets for access, equity and quality of learning outcomes for youth within a global post-2015 framework.

The extent to which this outcome and impact can be observed in 2020 is discussed in the following sections.

8.3.1. Outcome of PISA-D

The countries participating in PISA-D have confirmed the relevance and usefulness of the enhanced PISA instruments through their participation in the pilot. Moreover, an anonymous stakeholder survey undertaken in early 2018 by PISA-D development partner, the United Kingdom's Department for International Development (DFID), saw several country respondents anticipate that they will use the results of the PISA-D assessments to review curricula; update classroom assessment manuals and workbooks to align with the PISA-D cognitive framework; inform discussions with their Ministry of Education; and provide input into scheduled educational reform.

There are far more LMIC (**43 middle-income countries**) participating in PISA 2022 compared with the **PISA 2015 baseline of 26**. This is evidence that LMIC are willing and able to use the enhanced instruments produced by PISA-D. It is important to note that in all cases, these countries participation is not contingent upon the availability of external funding. While several LMICs are participating in PISA 2022 with the support of development partners, this external funding has been secured fir the duration of the PISA 2022 cycle. Nonetheless, it is clear that low income countries (LIC) are dependent on external funding to be able to participate in PISA and other international large-scale assessments and unless sustainable funding mechanisms are established to support these countries in particular, it is unlikely that they will be able to participate.

With the launching of 8 PISA-D national reports between December 2018 and March 2019 and the launch of 7 PISA 2018 national reports, based on the PISA-D model, in December 2019, there is clear evidence of how PISA data are being used by national Ministries of Education in LMIC to inform policy discussion and decision-making. It should be expected that in the coming months and years the national improvement plans of these 15 countries will include measures of learning outcomes for youth at the secondary education level. It is also reasonable to expect in these 15 countries that relevant links will be established between education targets at the primary and secondary levels and these will be related to the quality of learning.

It is also clear from the experience of PISA-D and the increased take up of participation in PISA by LMIC in PISA 2022, that the three risks to this project outcome identified at the outset of the initiative were managed effectively. The first of these risks was that capacity to implement PISA may be lacking and considered "too costly" to support in LMIC. The PISA-D experience shows that provided new countries are prepared effectively, the capacity to implement PISA successfully can be developed and supported in a cost-effective way – the cost of capacity building for the countries receiving this support in PISA 2022 is a total of Euros 210,000 per country for the duration of the cycle; i.e. four years.

The second risk identified to this project outcome was that there were perceptions of PISA as an assessment geared for OECD countries and relevant only for OECD economies. Again, the PISA-D experience and the increased take-up of participation in PISA by LMIC suggests that this risk too has been managed effectively. A third risk to this project outcome was a perception that the value of PISA participation is "ranking" in the PISA "league tables". This risk, too, has been managed effectively and the collaboration with the OECD over the national reports for 8 PISA-D participating countries and 7 PISA 2018 participating countries during 2018 and 2019 indicate that while countries are interested in international comparisons of performance, they regard the value of PISA participation to be much more than simply seeing themselves in the "ranking" in the PISA "league tables".

8.3.2. Impact of PISA-D

The main risk to the impact of PISA-D that was identified at the outset of the project's activities was that the focus of national and international goals and targets would not include learning outcomes or that these are only considered for primary education. At the design stage of PISA-D in 2013, it was not guaranteed that the post-2015 framework would include goals related to quality of learning outcomes. It is therefore pleasing to note that the SDG 4 framework includes:

- 5 targets that are related to learning outcomes (4.1, 4.2, 4.4, 4.6 and 4.7)

- 3 global indicators that measure learning outcomes

- 4 thematic indicators that measure learning outcomes

As noted in Chapter 7, PISA and PISA-D are being used to monitor progress towards the achievement of SDG Target 4.1. Data from PISA and PISA-D are currently the primary source for the UN system's monitoring of global indicator 4.1.1.c. Therefore, the main risk to the impact of PISA-D has been managed effectively. In fact, the main evidence of the PISA-D project's impact is that PISA, PISA-D and other internationally comparable assessments are being used to measure progress towards global learning goals in the post-2015 framework.

PISA results are certainly informing national education debates in the LMIC that participated in PISA-D following the launches of national reports in these during December 2018. It is also clear that PISA 2018 results are informing national debates in countries in Latin America, North Africa and Asia since December 2019. It is too early to judge whether the findings from PISA-D will contribute to reforms aimed at improving the quality of learning in the participating countries, but there is every reason to be hopeful that they will.

The SDG 4 architecture has facilitated increased numbers of LMIC participating in international discussions regarding how to improve learning outcomes, benchmarked over time and relative to other countries in a global framework. For example, the

Global Education Meeting (GEM) held in Brussels in December 2018 included several such discussions. This is leading to the situation where more LMIC are willing to include quality of learning outcomes as part of their educational priorities.

It was assumed that adequate financing would be available from LMICs' own resources and those of development partners to sustain their participation in PISA-D and subsequently PISA over time. This has largely proven to be the case, but financing does remain a substantial challenge to LMIC generally and low-income countries' (LIC) participation in particular in PISA and other international large-scale assessments over time. As part of its contribution to the SDG 4 monitoring, the OECD has collaborated closely with UNESCO, the UNESCO Institute for Statistics (UIS) and the Global Education Monitoring (GEM) Report team to make the investment case for monitoring the Education SDG. The four organisations estimate that USD$ 280 million per year is needed to monitor SDG 4 effectively against which only USD$ 148 million per year is spent currently and the majority of this is used to support monitoring in upper-middle- and-high- income-countries.

UNESCO, UIS, GEM Report team and the OECD conclude that to meet the shortfall on SDG 4 monitoring in LMIC of USD$ 132 million per year, it will be necessary to increase the domestic and external resources allocated to SDG 4 monitoring, especially to support learning assessments such as PISA, as well as national and regional assessments and household (multi-purpose) surveys. There is no doubt that if this financing gap is not closed, there will continue to be funding constraints to low-income-country participation in PISA and, more generally, continuing learning assessment data gaps for LMIC especially.

A second risk to the impact of PISA-D that was identified at the outset of the project's activities, was that political economy factors could prove a more challenging barrier to greater participation of LMIC in PISA (e.g. armed conflicts, fragility, natural hazards). These factors have not proven to be a major hindrance to increasing participation in PISA to date. However, there is no doubt that these factors could prove a more challenging barrier to greater participation of the poorest countries in PISA.

8.3.3. Independent review of PISA-D

At the first meeting of the IAG it was agreed that an independent review of PISA-D should be undertaken after the completion of the initiative. Terms of Reference for an independent review of PISA-D were developed by the OECD Secretariat and discussed and agreed by the IAG. The results of the review are expected to provide an assessment of the value and influence of the project's contribution to the education sectors of the participating countries, notably in building capacity for large-scale student assessment, their actual and anticipated use of the results and findings from the project to inform policy discussions across the education sector and reforms for improving educational outcomes, and the likely impact of the project on developing country participation in international assessments in the future, particularly PISA. The review is also intended to provide forward looking recommendations about future initiatives to support large-scale student learning assessment in developing countries and the post-project governance arrangements for PISA.

Unfortunately, as of late 2020, the OECD has been unable to raise sufficient funds (Euros 100 000 is required) from development partners to cover the costs of an independent review of PISA-D. The OECD will continue to look for the necessary funds to complete this activity and, if successful with fund raising, will manage the independent review in accordance with the Terms of reference agreed by the IAG.

8.4. Next steps

The OECD is seeking to build on the success of its PISA-D initiative and its experience of working in education with more than 40 LMIC Partners in PISA to continue to make the programme more accessible and relevant to a wider range of countries. In order to cope effectively with increased demand for participation, especially LMIC, the OECD has put in place a plan for managing a larger PISA programme. This plan includes:

- Optional support for new Partners to help them prepare for their participation in PISA and assist them with survey implementation.

- Encouragement for new Partners to concentrate in their first cycle on the implementation of PISA's core components and will not be offered the alternative to take on additional options such as optional questionnaires or tests, over-sampling and the optional out-of-school component until they have the experience of completing at least one cycle successfully.

- Promotion of peer-to-peer learning partnerships between a peer (an experienced PISA country) and a new country is to help new National Centres to benefit fully from their participation in PISA and to complete the necessary phases of project implementation.

- Enhancement of the way results are presented in the PISA international reports to reflect the wider participation in the programme and to make available the full range of data.

- Support for analysis and reporting has been offered to and taken up by 7 Partners as an option in PISA 2018 and 6, so far, in PISA 2022 and will also be offered for PISA 2025.

- Support for groups of countries to prepare regional reports is being discussed.

- The paper based assessment will be maintained for PISA 2022 and PISA 2025 and the comparability of paper- and computer-based scales will be investigated during scaling and discussed by the PISA TAG and reflected in the PISA international reports.

- Technical review of the PISA scaling methods to accommodate further non- Member participation.

- Continued enhancements are being incorporated into the PISA instruments to better describe the performance and contexts of a wider range of students.

- The contextual questionnaires in PISA 2022 will seek to extend PISA's measure of economic, social and cultural status (ESCS) to better describe the contexts of a broad variety of socio-economic contexts in all Partners and Member countries, and will also incorporate other variables to improve understanding of factors associated with performance in LMIC contexts.

- The experience of PISA-D with the out-of-school component will be further reviewed and approaches for an optional module to be integrated in PISA in the future will be taken forward. Currently, three options are being considered for this: (i) a *stand-alone* international option as part of a future PISA cycle (10 minutes core module and 35 minutes test with results linked to the PISA scale); (ii) an international option as part of a future PISA cycle *linked to a household survey*

(10 minutes core module and 35 minutes test with results linked to the PISA scale); and (iii) a *PISA module (a 15-20 minutes test based on the PISA-D Strand C test) integrated with a multi-topic household survey* designed solely to discriminate whether respondent is above or below 406 points on reading and 419 points on mathematics – the SDG 4 benchmarks for minimum levels of proficiency – may be part of a future PISA cycle or a completely separate study.

- Streamlining of PISA meetings of NPMs in light of the increased number of participants – these streamlining measures include limiting the number of country representatives at each meeting; adding content-specific days to NPM meetings; holding special NPM meetings for new countries; improving NPM governance; and improving security at NPM meetings.

In addition, with the support of its development partners, the OECD proposes to take to scale in PISA approaches, methods and tools that solve persistent educational challenges, particularly exclusion and equity issues.

If the OECD is successful in raising funds for an independent review of PISA-D, the results of this exercise will be incorporated into all of the next steps outlined above.

The OECD will continue to support capacity development and knowledge exchange among LMIC participants in PISA with activities that strengthen national capacity for learning assessment through peer review and exchange; creation of learning modules and diagnostic tools; and face-to-face exchange.

The OECD also proposes activities that aim to consolidate, where appropriate, and extend, where this is possible, knowledge in LMIC about how to improve educational outcomes and strengthen national education systems.

References

Adams, R. and J. Cresswell (2016), "PISA For Development Technical Strand A: Enhancement of PISA Cognitive Instruments", *OECD Education Working Papers*, No. 126, OECD Publishing, Paris, https://doi.org/10.1787/5jm5fb3f85s0-en.

Bloem, S. (2013), "PISA in Low and Middle Income Countries", *OECD Education Working Papers*, No. 93, OECD Publishing, Paris, https://doi.org/10.1787/5k41tm2gx2vd-en.

Carr-Hill, R. (2015), "PISA for Development Technical Strand C: Incorporating out-of-school 15- year-olds in the assessment", *OECD Education Working Papers*, No. 120, OECD Publishing, Paris, http://dx.doi.org/10.1787/5js0bsln9mg2-en.

Cresswell, J., U. Schwantner and C.Waters (2015), *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data*, PISA, The World Bank, Washington, D.C./OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264248373-en.

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

OECD (2013a), "What Makes Urban Schools Different?", *PISA in Focus*, No. 28, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k46l8w342jc-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 (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 (2016), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*, PISA, OECD Publishing, Paris, DOI: http://dx.doi.org/10.1787/9789264255425-en

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

OECD (2016b), *PISA 2015 Database*, http://www.oecd.org/pisa/data/2015database/ (accessed on 03 July 2018)

OECD (2017), "Where did equity in education improve over the past decade?", *PISA in Focus*, No. 68, OECD Publishing, Paris, http://dx.doi.org/10.1787/33602e45-en.

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

OECD (2018a), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris,https://doi.org/10.1787/eag-2018-en.

Spaull, N. (2017), "Who Makes It Into PISA?: Understanding the Impact of PISA Sample Eligibility Using Turkey as a Case Study (PISA 2003 - PISA 2012)", *OECD Education Working Papers*, No. 154, OECD Publishing, Paris, http://dx.doi.org/10.1787/41d175fc-en.

UNICEF and UNESCO Institute for Statistics (2010), All in School: Global Initiative in Out-of-School Children, http://allinschool.org/http://www.uis.unesco.org/Education/Documents/oosci-global-report-en.pdf.

UNESCO (2015), Education 2030 Incheon Declaration and Framework for Action: Towards inclusive and equitable quality education and lifelong learning for all, available at: www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/ED/pdf/FFA_Complet_Web-ENG.pdf

UNESCO Institute of Statistics (2015), *Fixing the Broken Promise of Education for All: Findings from the Global Initiative on Out-of-School Children*, UNESCO-UIS, Montreal, www.uis.unesco.org/Education/Documents/oosci-global-report-en.pdf

UNESCO Institute of Statistics (2016), *Making Education Count for Development: Data Collection and Availability in Six PISA for Development Countries*, PISA, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264255449-en

Willms, J.D. (2015), *Educational Prosperity*, Fredericton: The Learning Bar Inc

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. DOI: http://dx.doi.org/10.1787/5js1kv8crsjf-en.

World Bank (2017), *World Bank Country and Lending Groups*, https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Annex A. Logical Framework

Narrative Summary	Objectively Verifiable Indicators – OVIs	Means of Verification -MOVs	External Factors *(Positive Assumptions)*	External Factors *(Risks)*
IMPACT The majority of countries – high income and LMIC - are able to effectively monitor progress towards national and international goals and targets for access, equity and quality of learning outcomes for youth within a global post-2015 framework.	- PISA and other internationally comparable assessments used to measure progress towards global learning goals in the post-2015 framework. - PISA results inform national education debates in LMIC and findings contribute to reforms aimed at improving the quality of learning. - Increased numbers of LMIC participate in international discussions regarding how to improve learning outcomes, benchmarked over time and relative to other countries in a global framework	- PISA main survey reports for 2021 and beyond. - National and multi-national reports developed by PISA participating countries and development partners.	- LMIC countries are willing to include quality of learning outcomes as part of their educational priorities. - The post-2015 framework includes goals related to quality of learning outcomes. - Adequate financing is available from LMIC' own resources and those of development partners to sustain their participation in PISA over time.	- The focus of national and international goals and targets does not include learning outcomes or these are only considered for primary education. - Political economy factors prove a more challenging barrier to greater participation of LMIC in PISA (e.g. armed conflicts, fragility, natural hazards).
OUTCOME Increased numbers of developing countries use PISA assessments from 2021 onwards to monitor progress towards national improvement targets, to comparatively analyse factors associated with student outcomes, for institutional capacity-building and for tracking international education targets within a post-2015 framework.	- Countries confirm the relevance and usefulness of the enhanced PISA instruments through their participation in the pilot. - More LMIC participate in PISA cycles from 2021 onwards (compared with 2015 baseline). - Numbers of national improvement plans that include measures of learning outcomes for youth at the secondary education level.	- Proceedings from the International seminar with participating project countries. - Technical reports and results from pilot. - National and multi-national reports developed by PISA participating countries and development partners.	- LMIC countries are willing and able to use the enhanced instruments. - Relevant links established by countries between education targets at the primary and secondary levels and related to quality of learning .	- Capacity to implement PISA may be lacking and considered "too costly" to support in LMIC. - Perceptions of PISA as an assessment geared for OECD countries and relevant only for OECD economies. - Perception that the value of PISA participation is "ranking" in the PISA "league tables".

Narrative Summary	Objectively Verifiable Indicators – OVIs	Means of Verification -MOVs	External Factors (Positive Assumptions)	External Factors (Risks)
OUTPUTS (RESULTS) 1. Contextual questionnaires and data-collection instruments enhanced (system-level, for students, parents and schools) 2. Descriptive power of cognitive assessments enhanced (reading, mathematics and science) 3. An analytical framework and methodological approach for including out-of-school 15-year-olds in assessments developed. 4. Country capacity in assessment and analysis strengthened among participants. 5. Engagement established with LMIC and partners for peer-to-peer analysis and learning opportunities to support the UN-led post-2015 process.	- Questionnaires and data collection instruments developed and trialled on schedule. - Item parameters, inter-coder reliability and scaling outputs comparable to main PISA studies. - Approaches to out-of-school children developed and trialled according to schedule. - Training in and support for assessment and analysis delivered to project country personnel. - Outreach activities, including meetings, workshops and an international seminar (Y3) implemented.	- PISA for Development final report (narrative and technical). - Summary records and proceedings of meetings, technical workshops and international seminar. - Quarterly project statistical and narrative reports. - Quarterly supervision reports of project managers and contractors. - Final report on results at the end of Y3.	- Ownership of process by national governments of participating countries and institutional support provided for implementation. - Steering Group and Expert Group functions effectively. - Institutional and technical partnerships established for deep expertise in thematic areas (e.g. with UNICEF and UIS for out-of-school youth). - PISA and results from pilot are considered in the UN-led discussions of the post-2015 process.	- Sufficient institutional support is not provided by national governments and bureaucracies for efficient and effective implementation. - Significant parts of the population in participating countries, particularly among the most excluded and vulnerable, are not included in the project. - Some of the constructs used in PISA cannot be adequately applied, adapted or operationalised in some LMIC contexts to ensure international comparability.

MAIN ACTIVITIES

1. International Advisory Group formed to provide technical and implementation guidance – Y1 (2014)
2. Institutional and technical partnerships established – Y1
3. Expert group formed to oversee implementation: Y1
4. International contractors awarded contracts:Y1-Y2 (2015)
5. Coordination and preparation for implementation between OECD, development partners and project countries: Y1
6. Instruments designed, developed, translated/adapted, field-trialled, equated and validated: Y2-Y3 (2015-2016)
7. Main Survey data collection: Y4 (2017) and Y5 (2018)
8. Oversight, monitoring and coordination visits and meetings: Y1-Y6
9. Engagement and contribution to UN-led post-2015 framework discussions regarding the potential role of PISA: Y1-Y6 (2019).
10. Analysis and reporting of results to participating countries: end of Y5 and end of Y6

INPUTS (Budget): 10,526,642

A. Technical Development, Implementation, Analysis and Reporting: 9,946,916
A1. International Advisory Group: 375,766
A2. Instrument development, field trials, local assessment implementation and related services: 6 949 371
A3. Technical oversight, coordination, analysis and reporting: 2 311 329
A4. Engagement for peer-to-peer learning and contribution to UN-led post-2015 process: 310 450
Overhead: 579 726

B. In-country Operations per participating country (incurred by Ministry): xxx,000*
*This will vary from country to country and arrangements will be between development partners and government ministries.