Introduction

OVERVIEW

Parents, students, teachers, governments and the general public – all stakeholders – need to know how well their education systems prepare students for real-life situations. Many countries monitor students' learning to evaluate this. Comparative international assessments can extend and enrich the national picture by providing a larger context within which to interpret national performance. They can show what is possible in education, in terms of the quality of educational outcomes as well as in terms of equity in the distribution of learning opportunities. They can support policy targets by establishing measurable goals achieved by other systems and help to build trajectories for reform. They can also help countries to work out their relative strengths and weaknesses and monitor progress.

In response to the need for cross-nationally comparable evidence on student performance, the Organisation for Economic Co-operation and Development (OECD) launched the Programme for International Student Assessment (PISA) in 1997. PISA represents a commitment by governments to monitor the outcomes of education systems by measuring student achievement on a regular basis and within an internationally agreed common framework. It aims to provide a new basis for policy dialogue and for collaboration in defining and implementing educational goals, in innovative ways that reflect judgments about the skills that are relevant to adult life.

PISA is a collaborative effort undertaken by its participants – the OECD member countries as well as over 30 non-member partner countries and economies – to measure how well students, at age 15, are prepared to meet the challenges they may encounter in future life. Age 15 is chosen because at this age, students are approaching the end of compulsory education in most OECD countries. PISA, jointly guided by the participating governments, brings together the policy interests of countries with scientific expertise at both national and international levels. PISA has been measuring the knowledge, skills and attitudes of 15-year-olds over the last twelve years and is therefore able to give some insight into how countries are faring over time.

The PISA assessment takes a broad approach to measuring knowledge, skills and attitudes that reflect current changes in school priorities, moving beyond the school-based approach towards the use of knowledge in tasks and challenges likely to be encountered in home and work life outside school. It is based on a dynamic model of lifelong learning in which new knowledge and skills necessary for successful adaptation to a changing world are continuously acquired throughout life. PISA focuses on competencies that 15-year-old students will need in the future and seeks to assess what they can do with what they have learnt – reflecting the ability of students to continue learning throughout their lives by applying what they learn in school to non-school environments, evaluating their choices and making decisions. The assessment is informed, but not constrained, by the common denominator of national curricula. Thus, while it does assess students' knowledge, PISA also examines their ability to reflect, and to apply their knowledge and experience to real-life issues in a reflective way. For example, in order to understand and evaluate scientific advice on food safety, an adult would need not only to know some basic facts about the composition of nutrients, but also to be able to apply that information. The term "literacy" is used to encapsulate this broader concept of knowledge and skills, and the PISA assessment aims to determine the extent to which 15-year-old students can activate various cognitive processes that would enable them to

make effective use of the reading, mathematical and scientific knowledge and skills they have acquired throughout their schooling and related learning experiences up to that point.

PISA is designed to collect information through three-yearly assessments and presents data on domain-specific knowledge and skills in reading, mathematics and science of students, schools and countries. It combines the assessment of reading, mathematics and science with information on students' home background, their approaches to learning, their learning environments and their familiarity with computers. Thereby, PISA provides insights into the factors that influence the development of skills and attitudes at home and at school, and examines how these factors interact and what the implications are for policy development.

PISA uses: *i*) strong quality assurance mechanisms for translation, sampling and test administration; *ii*) measures to achieve cultural and linguistic breadth in the assessment materials, particularly through countries' participation in the development and revision processes for the production of the items; and *iii*) state-of-the-art technology and methodology for data handling. The combination of these measures produces high quality instruments and outcomes with superior levels of validity and reliability to improve the understanding of education systems as well as students' knowledge, skills and attitudes.

This publication presents the theory underlying the PISA 2012 assessment, including a re-developed and expanded framework for mathematical literacy, incorporating processes in which students engage when they solve problems as a new reporting dimension. It includes also a new optional computer-based assessment of mathematics (CBAM), reflecting the importance of Information and Communication Technologies (ICTs) for working mathematically in modern societies. It also provides the basis for the assessment of reading and science. Within each domain, the knowledge content that students need to acquire is outlined, as well as the processes that need to be performed and the contexts in which knowledge and skills are applied. It also illustrates the domains and their aspects with sample tasks. Finally, the theory underlying the context questionnaires is presented. The questionnaires are used to gather information from students, schools and parents on the students' home background and attitudes, their learning histories and their learning environments at school.

BASIC FEATURES OF PISA 2012

PISA 2012 is the fifth cycle of a data strategy defined in 1997 by participating countries. The OECD publications *Measuring Student Knowledge and Skills – A New Framework for Assessment* (1999), *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (2003), *Assessing Scientific, Reading and Mathematical Literacy – A Framework for PISA 2006* (2006) and *PISA 2009 Assessment Framework – Key competencies in Reading, Mathematics and Science* (2009) presented the conceptual framework underlying the first four cycles of PISA. The results from those cycles were presented in the OECD publications *Knowledge and Skills for Life – First Results from PISA 2000* (2001), *Learning for Tomorrow's World: First Results from PISA 2003* (2004), *PISA 2006: Science Competencies for Tomorrow's World* (2007) and *PISA 2009 Results – Volumes I to VI* (2010). All publications are also available on the PISA website: *www.pisa.oecd.org*. The results allow national policy makers to compare the performance of their education systems with those of other countries. Similar to the previous assessments, the 2012 assessment covers reading, mathematics and science, with the major focus on mathematical literacy. Students also respond to a background questionnaire, and additional supporting information is gathered from the school authorities. In 11 countries and economies information is also gathered from the students' parents. Sixty-six countries and economies, including all 34 OECD member countries, are taking part in the PISA 2012 assessment.

Since the aim of PISA is to assess the cumulative yield of education systems at an age where compulsory schooling is still largely universal, testing focuses on 15-year-olds enrolled in both school-based and work-based educational programmes. Between 4 500 and 10 000 students from at least 150 schools are typically tested in each country, providing a good sampling base from which to break down the results according to a range of student characteristics.

The primary aim of the PISA assessment is to determine the extent to which young people have acquired the wider knowledge and skills in reading, mathematics and science that they will need in adult life. The assessment of cross-curricular competencies continues to be an integral part of PISA 2012. The main reasons for this broadly oriented approach are:

- Although specific knowledge acquisition is important in school learning, the application of that knowledge in adult life depends crucially on the acquisition of broader concepts and skills. In reading, the capacity to develop interpretations of written material and to reflect on the content and qualities of text are central skills. In mathematics, the ability

to answer familiar textbook questions must be supplemented by being able to reason quantitatively, to represent relationships or dependencies, and to connect the context and structure of a problem with mathematics when it comes to deploying mathematical skills in real world problems. In science, having specific knowledge, such as the names of plants and animals, is of less value than understanding broad topics such as energy consumption, biodiversity and human health in thinking about the issues under debate in the adult community.

- In an international setting, a focus on curriculum content would restrict attention to curriculum elements common to all or most countries. This would force many compromises and result in an assessment too narrow to be of value for governments wishing to learn about the strengths and innovations in the education systems of other countries.
- Certain broad, general skills are essential for students to develop. They include communication, adaptability, flexibility, problem solving and the use of information technologies. These skills are developed across the curriculum and an assessment of them requires a broad cross-curricular focus.

Box 0.1 **What is PISA?**

Basics

- An internationally standardised assessment that was jointly developed by participating countries and administered to 15-year-olds in educational programmes.
- A survey implemented in 43 countries and economies in the first cycle (32 in 2000 and 11 in 2002), 41 in the second cycle (2003), 57 in the third cycle (2006) and 75 in the fourth cycle (65 in 2009 and 10 in 2010). In PISA 2012, 66 countries and economies participated.
- The test is typically administered to between 4 500 and 10 000 students in each country/economy.

Content

- PISA 2012 covers the domains of mathematics, reading and science not only in terms of whether students can reproduce specific subject matter knowledge, but also whether they can extrapolate from what they have learnt and apply their knowledge in novel situations. Two other domains were included in the PISA 2012 cycle: problem solving, in which not all countries participated because of technical issues, and financial literacy, which was administered as an option by some countries.
- Emphasis is on the mastery of processes, the understanding of concepts and the ability to function in various situations within each domain.

Methods

- Paper-and-pencil tests are used, with assessments lasting a total of two hours for each student.

In a range of countries and economies, an additional 40 minutes are devoted to the computer-based assessment of mathematics and reading.

- Test items are a mixture of multiple-choice items and questions requiring students to construct their own responses. The items are organised in groups based on a passage setting out a real-life situation.
- A total of about 390 minutes of test items is covered, with different students taking different combinations of test items.
- Students answer a background questionnaire, which takes 30 minutes to complete, providing information about themselves and their homes. School principals are given a 20-minute questionnaire about their schools. In some countries and economies, optional short questionnaires are administered to: *i*) parents to provide further information on past and present reading engagement at the students' homes; and *ii*) students to provide information on their access to and use of computers as well as their educational history and aspirations.

Assessment cycle

- The assessment takes place every three years with a strategic plan in place extending through to 2015.
- Each of these cycles looks in depth at a major domain, to which two-thirds of testing time is devoted; the other domains provide a summary profile of skills. Major domains have been reading in 2000 and 2009, mathematics in 2003 and science in 2006. In 2012, the major domain is again mathematical literacy.

Outcomes

- A basic profile of knowledge and skills among 15-year-old students.
- Contextual indicators relating results to student and school characteristics. Trend indicators showing how results change over time.
- A valuable knowledge base for policy analysis and research.

PISA is not a single cross-national assessment of the reading, mathematics and science skills of 15-year-old students. It is an ongoing programme that, over the longer term, will lead to the development of a body of information for monitoring trends in the knowledge and skills of students in various countries as well as in different demographic subgroups of each country. On each occasion, one domain is tested in detail, taking up nearly two-thirds of the total testing time. This data collection strategy provides a thorough analysis of achievement in each area every nine years and a trend analysis every three. The major domain was reading in 2000 and 2009, mathematics in 2003 and science in 2006. In 2012, it is mathematics again, building on a modified mathematics framework which incorporates the computer-based assessment of mathematics and includes the mathematical processes which students undertake when using mathematical literacy and the fundamental mathematical capabilities which underlie those processes (see Chapter 1). The reading and science frameworks for PISA 2012 are the same as for the previous assessment (see Chapters 2 and 3, respectively).

Similar to previous PISA cycles, the paper-and-pen assessment was designed as a two-hour test comprising four 30-minute clusters of test material from one or more cognitive domains. Information was obtained from about 390 minutes worth of test items. For each country, the total set of questions was packaged into 13 linked test booklets. Financial literacy, an option in the paper-and-pen assessment, was allocated two clusters (that is, 60 minutes of testing time) in the 2012 main survey. Each booklet was taken by a sufficient number of students for appropriate estimates to be made of the achievement levels on all items by students in each country and in relevant sub-groups within a country (such as boys and girls, and students from different social and economic contexts). Students also spent 30 minutes answering a background questionnaire. Applying a rotated design to the student questionnaire allowed for more material to be used in the study. Some questions were answered by all students, as in previous cycles, some by sub-samples of students.

In addition to this core assessment, 44 countries and economies participated in a computer-based assessment of problem solving, and among them, 32 participated in a computer-based assessment of reading and mathematics. The duration of the PISA 2012 computer-delivered assessment was 40 minutes. A total of 80 minutes of problem-solving material was organised into four 20-minute clusters. Students from countries not participating in the optional computer-based assessment of mathematics and digital reading did two of the clusters according to a balanced rotation design. Students from countries also participating in the optional computer-based assessment of mathematics and digital reading did two, one or none of the four problem-solving clusters according to a separate balanced rotation design. The optional computer-based component contained a total of 80 minutes of mathematics material and 80 minutes of reading material. The material for each domain was arranged in four clusters of items, with each cluster representing 20 minutes of testing time. All material for computer delivery was arranged in a number of rotated test forms, with each form containing two clusters. Each student did one form, representing a total testing time of 40 minutes.

The PISA assessment provides three main types of outcomes:

- Basic indicators that provide a baseline profile of the knowledge and skills of students.
- Indicators derived from the contextual questionnaire that show how such skills relate to important demographic, social, economic and educational variables.
- Indicators on trends that emerge from the on-going nature of the data collection and that show changes in outcome levels and distributions, and in relationships between student-level and school-level background variables and outcomes.

Although indicators are an adequate means of drawing attention to important issues, they do not provide answers to policy questions. Therefore, PISA has also developed a policy-oriented analysis plan that goes beyond the reporting of indicators.

WHAT MAKES PISA UNIQUE

PISA focuses on young people's ability to use their knowledge and skills to meet real-life challenges. This orientation reflects a change in the goals and objectives of curricula themselves, which are increasingly concerned with what students can do with what they learn at school and not only with whether they have mastered specific curricular content.

Key features driving the development of PISA have been its:

- Policy orientation, which connects data on student learning outcomes with data on students' characteristics and on key factors shaping their learning inside and outside school in order to draw attention to differences in performance patterns and to identify the characteristics of schools and education systems that have high performance standards.
- Innovative "literacy" concept, which is concerned with the capacity of students to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they pose, solve and interpret problems in a variety of situations.

- Relevance to lifelong learning, which does not limit PISA to assessing students' curricular and cross-curricular competencies, but also asks them to report on their own motivation to learn, their beliefs about themselves and their learning strategies.
- Regularity, which enables countries to monitor their progress in meeting key learning objectives.
- Breadth of geographical coverage and collaborative nature, which in PISA 2012 encompasses the 34 OECD member countries and over 30 partner countries and economies.

The relevance of the knowledge and skills measured by PISA is confirmed by recent studies tracking young people in the years after they have been assessed by PISA. Studies in Australia, Canada and Denmark display a strong relationship between the performance in reading on the PISA 2000 assessment at age 15 and the chance of a student completing secondary school and of carrying on with post-secondary studies at age 19. For example, Canadian students who had achieved reading proficiency Level 5 at age 15 were 16 times more likely to be enrolled in post-secondary studies when they were 19 years old than those who had not reached the reading proficiency Level 1.

PISA is the most comprehensive and rigorous international programme to assess student performance and to collect data on the student, family and institutional factors that can help to explain differences in performance. Decisions about the scope and nature of the assessments and the background information to be collected are made by leading experts in participating countries, and are steered jointly by governments on the basis of shared, policy-driven interests. Substantial efforts and resources are devoted to achieving cultural and linguistic breadth and balance in the assessment materials. Stringent quality assurance mechanisms are applied in translation, sampling and data collection. As a consequence, the results of PISA have a high degree of validity and reliability, and can significantly improve understanding of the outcomes of education in the world's economically most developed countries, as well as in a growing number of countries at earlier stages of economic development.

Across the world, policy makers are using PISA findings to: gauge the knowledge and skills of students in their own country in comparison with those of the other participating countries; establish benchmarks for educational improvement, for example, in terms of the mean scores achieved by other countries or their capacity to provide high levels of equity in educational outcomes and opportunities; and understand relative strengths and weaknesses of their education systems. The interest in PISA is illustrated by the many reports produced in participating countries, the numerous references to the results of PISA in public debates and the intense media attention shown to PISA throughout the world.

AN OVERVIEW OF WHAT IS BEING ASSESSED IN EACH DOMAIN

Box B presents a definition of the three domains assessed in PISA 2012. The definitions all emphasise functional knowledge and skills that allow one to participate actively 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 equipped to take part in decision-making processes. In the more complex tasks in PISA, students are asked to reflect on and evaluate material, not just to answer questions that have single correct answers. The definitions address the capacity of students to extrapolate from what they have learnt, and to apply their knowledge in novel settings. The definitions also focus on the students' capacity to analyse, reason and communicate effectively, as they pose, solve and interpret problems in a variety of situations.

Box 0.2 **Definitions of the domains**

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.

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.

Scientific literacy: An individual's scientific knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena, and to draw evidence-based conclusions about science-related issues, understanding of the characteristic features of science as a form of human knowledge and enquiry, awareness of how science and technology shape our material, intellectual, and cultural environments, and willingness to engage in science-related issues, and with the ideas of science, as a reflective citizen.

Mathematical literacy (elaborated in Chapter 1) is concerned with the ability of students to analyse, reason, and communicate ideas effectively as they pose, formulate, solve, and interpret solutions to mathematical problems in a variety of situations. The PISA mathematics assessment has, so far, been designed in relation to the:

- *Processes*: These are defined in terms of three categories (*formulating situations mathematically; employing mathematical concepts, facts, procedures and reasoning*; and *interpreting, apply and evaluating mathematical outcomes* – referred to in abbreviated form as *formulate, employ* and *interpret*) and describe what individuals do to connect the context of a problem with the mathematics and thus solve the problem. These three processes each draw on the seven fundamental mathematical capabilities (*communication; mathematising; representation; reasoning and argument; devising strategies for solving problems; using symbolic, formal and technical language and operations; using mathematical tools*) which in turn draw on the problem solver's detailed mathematical knowledge about individual topics.
- *Content*: This is defined mainly in terms of four overarching ideas (*quantity, space and shape, change and relationships,* and *uncertainty and data*) which relate to familiar curricular strands such as numbers, algebra and geometry in overlapping and complex ways.
- *Contexts*: This is defined in terms of the aspect of an individual's world in which the problems are placed. The framework identifies four categories: *personal, educational, societal* and *scientific*.

Reading literacy (elaborated in Chapter 2) is defined in terms of students' ability to understand, use and reflect on written text to achieve their purposes. In PISA, reading literacy is assessed in relation to the:

- *Text format:* PISA uses *continuous texts* or prose organised in sentences and paragraphs and in addition *non-continuous texts* that present information in other ways, such as in lists, forms, graphs, or diagrams. It has also distinguished between a range of prose forms, such as narration, exposition and argumentation.
- *Processes (aspects):* Students are not assessed on the most basic reading skills, as it is assumed that most 15-year-old students will have acquired these. Rather, they are expected to demonstrate their proficiency in *accessing and retrieving information, forming a broad general understanding of the text, interpreting it, reflecting on its contents and reflecting on its form and features*.
- *Situations:* These are defined by the use for which the text was constructed. For example, a novel, personal letter or biography is written for people's personal use; official documents or announcements for public use; a manual or report for occupational use; and a textbook or worksheet for educational use. Since some groups may perform better in one reading situation than in another, it is desirable to include a range of types of reading in the assessment items.

Scientific literacy (elaborated in Chapter 3) is defined as the ability to use scientific knowledge and processes not only to understand the natural world but to participate in decisions that affect it. The PISA science assessment is designed in relation to:

- *Scientific knowledge or concepts:* These constitute the links that aid understanding of related phenomena. In PISA, while the concepts are the familiar ones relating to *physics*, *chemistry*, *biological sciences* and *earth and space* sciences, they are applied to the content of the items and not just recalled.
- *Processes:* These are centred on the ability to acquire, interpret and act upon evidence. Three such processes present in PISA relate to: *describing, explaining and predicting scientific phenomena, understanding scientific investigation,* and *interpreting scientific evidence and conclusions*.
- *Contexts:* These concern the application of scientific knowledge and the use of scientific processes applied. The framework identifies three main areas: science in *life and health*, science in *Earth and environment*, and science in *technology*.

ASSESSING AND REPORTING PISA 2012

Similar to the previous assessments in PISA, the assessment in 2012 mainly consisted of pencil and paper instruments. In addition, a computerised assessment of reading of electronic texts was carried out in a range of countries and economies. Both the paper-and-pencil assessment and the computer-based assessment included a variety of types of questions. Some required students to select or produce simple responses that can be directly compared with a single correct answer, such as multiple-choice or closed-constructed response items. These questions had either a correct or incorrect answer and often assess lower-order skills. Others were more constructive, requiring students to develop their own responses designed to measure broader constructs than those captured by more traditional surveys, allowing for a wider range of acceptable responses and more complex marking that can include partially correct responses.

Not all students answered all questions in the assessment. For the core paper-and-pencil assessment of mathematics, reading and science, the PISA 2012 test units were arranged in clusters that are scheduled 30 minutes of assessment time. In all paper-based tests, the booklets include four clusters (except one booklet for students with special needs, which consist of two clusters).

For the assessment of mathematics, countries could implement one of the two alternative sets of booklets which were provided in PISA 2012. The first set included 13 booklets that comprised items distributed across a range of difficulty similar to that of previous cycles. These booklets included four clusters according to a rotated test design among the seven mathematics clusters, three reading clusters and three science clusters. There was at least one mathematics cluster in each booklet. The second set also contained items covering the full range of difficulty, but included more items at the easier end of the range, in order to obtain better descriptive information about what students at the lower end of the ability spectrum know, understand and can do as mathematical problem solvers. All participating countries and economies administered 11 common clusters: five clusters of mathematics items, three clusters of reading items and three clusters of science items. In addition, countries administered one of two alternative pairs of mathematics clusters. Regardless of countries' choice of cluster, the performance of students in all participating countries and economies is represented on a common mathematical literacy scale. For the countries that chose the financial literacy test, two additional booklets were designed, and one for the students with special needs.

For the countries and economies which participated in the computer-based assessment, the test forms included two clusters of 20 minutes each. In the countries and economies which only chose to test problem solving on computers, test forms comprising two clusters were administered according to a rotated design. The test material consisted of eight test forms with two clusters each, and every student taking part was given one of the eight test forms to work on. In the countries and economies which chose to test problem solving, mathematics and reading on computers, test forms comprising two clusters were administered. The test material consisted of 24 test forms with two clusters each (according to a rotated design from four problem-solving clusters, four mathematics clusters, and two reading clusters).

For the paper-and-pencil assessment as well as the computerised assessment, knowledge and skills were assessed through units consisting of a stimulus (e.g. text, table, chart, figures, etc.) followed by a number of tasks associated with this common stimulus. This is an important feature, allowing questions to go into greater depth than if each question were to introduce a wholly new context. It allows time for the student to digest material that can then be used to assess multiple aspects of performance.

Results from PISA have been reported using scales with an average score of 500 and a standard deviation of 100 for all three domains, which means that two-thirds of students across OECD countries scored between 400 and 600 points. These scores represent degrees of proficiency in a particular domain. Reading literacy was the major domain in 2000, and the reading scales were divided into five levels of knowledge and skills. The main advantage of this approach is that it is useful for describing what substantial numbers of students can do with tasks at different levels of difficulty. Additionally, results were also presented through three *aspect* subscales of reading: *accessing and retrieving* information, *integrating and interpreting* texts, and *reflecting and evaluating* texts. A proficiency scale was also available for mathematics and science, though without levels therefore recognising the limitation of the data from minor domains. PISA 2003 built upon this approach by specifying six proficiency levels for the mathematics scale, following a similar approach to what was done in reading. There were four *content* subscales in mathematics*: space and shape, change and relationships, quantity,* and *uncertainty.* In a similar manner, the reporting of *science* in PISA 2006 specified six proficiency levels for the science scale. The three *competency* subscales in science related to *identifying scientific issues*, *explaining phenomena scientifically* and *using scientific evidence*. Additionally, country performance was compared on the bases of *knowledge about science* and *knowledge of science*. The three main areas of knowledge of science were *physical systems*, *living systems* and *earth and space systems*.

PISA 2009 was the first time that reading literacy was re-assessed as a major domain, and provided trend results for all three domains of *reading*, *mathematics* and *science*. In PISA 2009, beyond Level 5, which was the highest described level of proficiency in reading in previous PISA reading assessments, a new Level 6 has been added to describe very high levels of reading proficiency. The previous bottom level of measured proficiency, Level 1, has been relabelled as Level 1a. A new level, Level 1b, describes students who would previously have been rated as "below Level 1", but who show proficiency in relation to a new set of tasks that is easier than those included in previous PISA assessments. These changes allow countries to know more about what kinds of tasks students with very high and very low reading proficiency are capable of. Apart from the additional levels, the meaning of being proficient at reading Levels 2, 3, 4 and 5 remains the same in PISA 2009 as in previous surveys.

In PISA 2012 mathematics was re-assessed as a major domain, and, in addition to the *content* subscales (with the *uncertainty* scale being re-named as *uncertainty and data* for improved clarity), three new subscales were developed to point to the three processes in which students as active problem solvers will engage. These three *process* subscales are *formulating situations mathematically*; *employing mathematical concepts, facts, procedures and reasoning;* and *interpreting, apply and evaluating mathematical outcomes*, abbreviated as *formulating*, *employing*, and *interpreting*.

THE CONTEXT QUESTIONNAIRES AND THEIR USE

To gather contextual information, PISA asks students and the principals of their schools to respond to questionnaires of around 30 minutes in length. These questionnaires are central to the analysis of results in terms of a range of student and school characteristics. Chapter 6 presents the questionnaire framework in detail. The questionnaires from all assessments (PISA 2000, 2003, 2006, 2009 and 2012) are available on the PISA website: *www.pisa.oecd.org.* The questionnaires seek information about:

- Students and their family backgrounds, including their economic, social and cultural capital.
- Aspects of students' lives, such as their attitudes towards learning, their habits and life inside school, and their family environment.
- Aspects of schools, such as the quality of the schools' human and material resources, public and private management and funding, decision-making processes, staffing practices and the school's curricular emphasis and extra-curricular activities offered.
- Context of instruction, including institutional structures and types, class size, classroom and school climate and reading activities in class.
- Aspects of learning and instruction in reading, including students' interest, motivation and engagement.

Three additional questionnaires are offered as international options:

- A *computer familiarity questionnaire* focusing on the availability and use of information and communications technology (ICT), including where ICT is mostly used, as well as on the students' ability to carry out computer tasks and their attitudes towards computer use.
- An *educational career questionnaire* collecting additional information on interruptions of schooling, on preparation for their future career, on support with language learning.
- A *parent questionnaire* focusing on a number of topics including the parents' perceptions of and involvement in their child's school, their support for learning in the home, school choice, their child's career expectation particularly in mathematics and their migration background.

The contextual information collected through the student and school questionnaires, as well as the optional computer familiarity, educational career and parent questionnaires, comprises only a part of the total amount of information available to PISA. Indicators describing the general structure of the education systems (their demographic and economic contexts – for example, costs, enrolments, school and teacher characteristics, and some classroom processes) and their effect on labour market outcomes are already routinely developed and applied by the OECD (e.g. the yearly OECD publication *Education at a Glance*).

COLLABORATIVE DEVELOPMENT OF PISA AND ITS ASSESSMENT FRAMEWORK

PISA represents a collaborative effort among the OECD member governments to provide an innovative kind of assessment of student achievement on a recurring basis. The assessments are developed co-operatively, agreed by participating countries, and implemented by national organisations. The constructive co-operation of students, teachers and principals in participating schools has been crucial to the success of PISA during all stages of the development and implementation.

The PISA Governing Board (PGB), representing all nations at the 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. This includes setting priorities for the development of indicators, for the establishment of the assessment instruments and for the reporting of the results. Experts from participating countries also serve on working groups charged with linking the PISA policy objectives with the best internationally available technical expertise in the different assessment domains. By participating in these expert groups, countries ensure that the instruments are internationally valid and take into account the cultural and educational contexts in OECD member countries. They also ensure that the assessment materials have strong measurement properties and that the instruments emphasise authenticity and educational validity.

Participating countries implement PISA at the national level, through National Project Managers (NPM), subject to the agreed administration procedures. National Project Managers play a vital role in ensuring that implementation is of high quality. They also verify and evaluate the survey results, analyses, reports and publications.

The design of the assessment of mathematics, reading, science, problem solving, financial literacy, and the design and development of questionnaires, as well as the implementation of the present survey, within the framework established by the PISA Governing Board, is the responsibility of an international consortium led by the Australian Council for Educational Research (ACER). Other partners or sub-contractors in this consortium include cApStAn Linguistic Quality Control and the Department of Experimental and Theoretical Pedagogy at the University of Liège (SPe) in Belgium, the Deutsches Institut für Pädagogische Forschung (DIPF) in Germany, the National Institute for Educational Policy Research (NIER) in Japan, WESTAT in the United States, the Educational Testing Service (ETS) in the United States, the Institutt for Lærerutdanning og Skoleutvikling (ILS) in Norway, Leibniz – Institute for Science Education (IPN) in Germany, and the TAO Initiative: CRP – Henri Tudor and Université de Luxembourg – EMACS in Luxembourg. The OECD Secretariat has overall managerial responsibility for the programme, monitors its implementation on a day-to-day basis, acts as the secretariat for the PGB, builds consensus among countries and serves as the interlocutor between the PGB and the international consortium charged with implementation. The OECD Secretariat is also responsible for the production of the indicators, and the analysis and preparation of the international reports and publications in co-operation with the international consortium and in close consultation with member countries both at the policy level (PGB) and at the implementation level (National Project Managers).

The development of the PISA frameworks has been a continuous effort since the programme was created in 1997 and can be described as a sequence:

- Development of a working definition for the assessment domain and description of the assumptions that underlie that definition.
- Evaluation of how to organise the tasks constructed in order to report to policy makers and researchers on student achievement in the domain, and identification of key characteristics that should be taken into account when constructing assessment tasks for international use.
- Operationalisation of key characteristics used in test construction, with definitions based on existing literature and experience in conducting other large-scale assessments.
- Validation of the variables and assessment of the contribution they each make to understanding task difficulty across the participating countries.
- Preparation of an interpretative scheme for the results.

While the main benefit of constructing and validating a framework for each of the domains is improved measurement, there are other potential benefits:

- A framework provides a common language and a vehicle for discussing the purpose of the assessment and what it is trying to measure. Such a discussion encourages the development of a consensus around the framework and the measurement goals.
- An analysis of the kinds of knowledge and skills associated with successful performance provides a basis for establishing standards or levels of proficiency. As the understanding of what is being measured and the ability to interpret scores along a particular scale evolve, an empirical basis for communicating a richer body of information to various constituencies can be developed.
- Identifying and understanding particular variables that underlie successful performance further the ability to evaluate what is being measured and to make changes to the assessment over time.
- The understanding of what is being measured and its connection to what we say about students provides an important link between public policy, assessment and research which, in turn, enhances the usefulness of the data collected.