PISA 2012 Assessment and Analytical Framework

MATHEMATICS, READING, SCIENCE, PROBLEM SOLVING AND FINANCIAL LITERACY

Programme for International Student Assessment

OECD

PISA

PISA 2012 Assessment and Analytical Framework

MATHEMATICS, READING, SCIENCE, PROBLEM SOLVING AND FINANCIAL LITERACY

OECD
BETTER POLICIES FOR BETTER LIVES

Please cite this publication as:
OECD (2013), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy,* OECD Publishing.
http://dx.doi.org/10.1787/9789264190511-en

ISBN 978-92-64-19052-8 (print)
ISBN 978-92-64-19051-1 (PDF)

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

Photo credits:
© khoa vu/Flickr/Getty Images
© Shutterstock/Kzenon
© Simon Jarratt/Corbis

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

Foreword

The OECD Programme for International Student Assessment (PISA), created in 1997, represents a commitment by the governments of OECD member countries to monitor the outcomes of education systems in terms of student achievement, within a common internationally agreed framework. PISA is a collaborative effort, bringing together scientific expertise from the participating countries and steered jointly by their governments on the basis of shared, policy-driven interests. Participating countries take responsibility for the project at the policy level. Experts from participating countries also serve on working groups that are charged with linking the PISA policy objectives with the best available substantive and technical expertise in the field of internationally comparative assessment. Through involvement in these expert groups, countries ensure that the PISA assessment instruments are internationally valid and take into account the cultural and curricular context of OECD member countries. They also have strong measurement properties, and place an emphasis on authenticity and educational validity. PISA 2012 represents a continuation of the data strategy adopted in 1997 by OECD countries. As in 2003, mathematical literacy is the focus of the PISA 2012 survey. The framework for assessing mathematics was fully revised for the PISA 2012 assessment and introduces three new mathematical processes that form the basis of developments in the reporting of PISA mathematics outcomes. A computer-based assessment of mathematics was also included in the 2012 cycle. The framework for assessing science was revised for PISA 2006 while the framework for assessing reading was revised for PISA 2009. Both of these frameworks remained unchanged in PISA 2012. The analytic framework that formed the basis of the development of the various questionnaire instruments was also redeveloped for PISA 2012.

Additions to the PISA 2012 assessment include a computer-based assessment of problem solving and an assessment of financial literacy. In 2003, problem solving became an assessment domain in PISA but was not reintroduced in the PISA 2006 and 2009 cycles. However, a new framework was devised for problem solving in PISA 2012 and additional assessment methodologies were implemented, allowing for the real-time capture of students' capabilities. In particular, the PISA 2012 assessment of problem solving was computer-based, and interactivity of the student with the problem is a central feature of the assessment. Financial literacy was included for the first time in the PISA assessment. Its framework provides a common language for discussion about financial literacy, a working definition of the domain, an articulated plan for developing items, and defines the relevant content, processes and contexts for the assessment of 15-year-old students in this domain.

This publication presents the guiding principles of the PISA 2012 assessment, which are described in terms of the skills students need to acquire, the processes that need to be performed and the settings in which knowledge and skills are applied. Further, it illustrates the assessment domains with a range of sample tasks.

The framework development for mathematics was undertaken jointly by the Australian Council for Educational Research (ACER), and Achieve, Inc., a USA-based educational development organisation. The framework development for all other cognitive domains, as well as the context questionnaire, was undertaken by the Australian Council for Educational Research.

The frameworks were developed by the expert panels, with the guidance of Raymond Adams, Barry McCrae, Petra Lietz, Juliette Mendelovits, Dara Ramalingam and Ross Turner from ACER. The mathematics expert group was chaired by Kaye Stacey from the University of Melbourne. The problem solving expert group was chaired by Joachim Funke from the University of Heidelberg. The reading expert group was chaired by Irwin Kirsch of Educational Testing Service in the United States of America. The science expert group was chaired by Rodger Bybee, formerly of the Biological Science Curriculum Study in the United States. The financial literacy expert group was chaired by Annamaria Lusardi of The George Washington University School of Business, in the United States of America. The questionnaire expert group was chaired by Eckhard Klieme of the German Institute for International Educational Research (DIPF) in Germany. The members of the expert groups are listed in Annex B of this publication. The frameworks have also been reviewed by expert panels in each of the participating countries. The chapters were drafted by the respective expert groups under the direction of their chairs. The publication was prepared by the OECD Secretariat, principally by Michael Davidson, Sophie Vayssettes, Pablo Zoido, Giannina Rech, Elisabeth Villoutreix, Marilyn Achiron and Elizabeth Del Bourgo.

The report is published on the responsibility of the Secretary-General of the OECD.

Table of Contents

BOXES

FIGURES

TABLES

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.

1

PISA 2012 Mathematics Framework

The PISA 2012 mathematics framework explains the theoretical underpinnings of the PISA mathematics assessment, including a new formal definition of mathematical literacy, the mathematical processes which students undertake when using mathematical literacy, and the fundamental mathematical capabilities which underlie those processes. The framework describes how mathematical content knowledge is organised into four content categories and outlines the content knowledge that is relevant to an assessment of 15-year-old students. It describes four categories of contexts in which students will face mathematical challenges. The framework specifies the proportions of items from each of the four content and context categories, each response format and each process, and describes the rotating booklet designs and questionnaires. Items of a range of difficulty are required. The optional computer-based assessment for mathematics is described, with discussion of the rationale and potential for future development. The categorisations are illustrated with seven units used in PISA surveys and field trials. Multiple quality control measures are described. The PISA assessment will measure how effectively countries are preparing students to use mathematics in every aspect of their personal, civic and professional lives, as part of their constructive, engaged and reflective citizenship.

INTRODUCTION

The assessment of mathematics has particular significance for PISA 2012, as mathematics is the major domain assessed. Although mathematics was assessed by PISA in 2000, 2003, 2006 and 2009, the domain was the main area of focus only in 2003.

The return of mathematics as the major domain in PISA 2012 provides the opportunity to make comparisons in student performance over time, but it also provides the opportunity to re-examine what is assessed in light of changes that have occurred in the field and in instructional policies and practices. An inherent challenge is developing an updated, state-of-the-art mathematics framework while retaining psychometric links to past mathematics assessments through a forward-looking and backward-looking trend line. The PISA 2012 framework is designed to make mathematics relevant to 15-year-old students more clear and explicit, while ensuring that the items developed remain set in meaningful and authentic contexts. The mathematical modelling cycle, used in earlier frameworks (e.g. OECD, 2003) to describe the stages individuals go through in solving contextualised problems, remains a key feature of the PISA 2012 framework. It is used to help define the mathematical processes in which students engage as they solve problems – processes that are being used for the first time in 2012 as a primary reporting dimension. A new optional computer-based assessment of mathematics (CBAM) is also available for countries in 2012.

The PISA 2012 mathematics framework is organised into several major sections. The first section, "Definition of mathematical literacy", explains the theoretical underpinnings of the PISA mathematics assessment, including the formal definition of the mathematical literacy construct. The second section, "Organising the domain", describes three aspects: *i*) the mathematical processes and the fundamental mathematical capabilities (in previous frameworks the "*competencies*") underlying those processes; *ii*) the way mathematical content knowledge is organised in the PISA 2012 framework, and the content knowledge that is relevant to an assessment of 15-year-old students (sub-scores are being reported for both the three mathematical process categories and the four mathematical content categories); *iii*) the contexts in which students will face mathematical challenges. The third section, "Assessing mathematical literacy", outlines structural issues about the assessment, including a test blueprint and other technical information. The several addenda include further descriptions of the fundamental mathematical capabilities, several illustrative PISA items and a reference list.

This framework was written under the guidance of the Mathematics Expert Group (MEG), a body appointed by the main PISA contractors with the approval of the PISA Governing Board (PGB). The ten MEG members include mathematicians, mathematics educators, and experts in assessment, technology, and education research from a range of countries. In addition, to secure more extensive input and review, a draft of the PISA 2012 mathematics framework was circulated for feedback to over 170 mathematics experts from over 40 countries. Achieve and the Australian Council for Educational Research (ACER), the two organisations contracted by the Organisation for Economic Co-operation and Development (OECD) to manage framework development, also conducted various research efforts to inform and support development work. Framework development and the PISA programme generally have been supported and informed by the ongoing work of participating countries (e.g. the research described in the 2010 OECD publication *Pathways to Success: How Knowledge and Skills at Age 15 Shape Future Lives in Canada*).

DEFINING MATHEMATICAL LITERACY

An understanding of mathematics is central to a young person's preparedness for life in modern society. A growing proportion of problems and situations encountered in daily life, including in professional contexts, require some level of understanding of mathematics, mathematical reasoning and mathematical tools, before they can be fully understood and addressed. Mathematics is a critical tool for young people as they confront issues and challenges in personal, occupational, societal, and scientific aspects of their lives. It is thus important to have an understanding of the degree to which young people emerging from school are adequately prepared to apply mathematics to understanding important issues and solving meaningful problems. An assessment at age 15 provides an early indication of how individuals may respond in later life to the diverse array of situations they will encounter that involve mathematics.

As the basis for an international assessment of 15-year-old students, it is reasonable to ask: "What is important for citizens to know and be able to do in situations that involve mathematics?" More specifically, what does competency in mathematics mean for a 15-year-old, who may be emerging from school or preparing to pursue more specialised training for a career or university admission? It is important that the construct of mathematical literacy, which is used in this report to denote the capacity of individuals to formulate, employ, and interpret mathematics in a variety of contexts, not be perceived as synonymous with minimal, or low-level, knowledge and skills. Rather, it is intended to describe

the capacities of individuals to reason mathematically and use mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. This conception of mathematical literacy supports the importance of students developing a strong understanding of concepts of pure mathematics and the benefits of being engaged in explorations in the abstract world of mathematics. The construct of mathematical literacy, as defined for PISA, strongly emphasises the need to develop students' capacity to use mathematics in context, and it is important that they have rich experiences in their mathematics classrooms to accomplish this. This is true for those 15-year-old students who are close to the end of their formal mathematics training, as well as those who will continue with the formal study of mathematics. In addition, it can be argued that for almost all students, the motivation to learn mathematics increases when they see the relevance of what they are learning to the world outside the classroom and to other subjects.

Mathematical literacy naturally transcends age boundaries. However, its assessment for 15-year-olds must take into account relevant characteristics of these students; hence, there is a need to identify age-appropriate content, language and contexts. This framework distinguishes between broad categories of content that are important to mathematical literacy for individuals generally, and the specific content topics that are appropriate for 15-year-old students. Mathematical literacy is not an attribute that an individual either has or does not have. Rather, mathematical literacy is an attribute that is on a continuum, with some individuals being more mathematically literate than others – and with the potential for growth always present.

For the purposes of PISA 2012, mathematical literacy is defined as follows:

> *Mathematical literacy is 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.*

Some explanatory remarks are provided below to highlight and clarify aspects of the definition that are particularly important.

A view of students as active problem solvers in PISA 2012

The focus of the language in the definition of mathematical literacy is on active engagement in mathematics, and is intended to encompass reasoning mathematically and using mathematical concepts, procedures, facts and tools in describing, explaining and predicting phenomena. In particular, the verbs 'formulate,' 'employ,' and 'interpret' point to the three processes in which students as active problem solvers will engage. *Formulating* mathematics involves identifying opportunities to apply and use mathematics – seeing that mathematics can be applied to understand or resolve a particular problem or challenge presented. It includes being able to take a situation as presented and transform it into a form amenable to mathematical treatment, providing mathematical structure and representations, identifying variables and making simplifying assumptions to help solve the problem or meet the challenge. *Employing* mathematics involves applying mathematical reasoning and using mathematical concepts, procedures, facts and tools to derive a mathematical solution. It includes performing calculations, manipulating algebraic expressions and equations or other mathematical models, analysing information in a mathematical manner from mathematical diagrams and graphs, developing mathematical descriptions and explanations and using mathematical tools to solve problems. *Interpreting* mathematics involves reflecting upon mathematical solutions or results and interpreting them in the context of a problem or challenge. It includes evaluating mathematical solutions or reasoning in relation to the context of the problem and determining whether the results are reasonable and make sense in the situation.

The language of the definition is also intended to integrate the notion of mathematical modelling, which has historically been a cornerstone of the PISA framework for mathematics (e.g. OECD, 2003), into the PISA 2012 definition of mathematical literacy. As individuals use mathematics and mathematical tools to solve problems in contexts, their work progresses through a series of stages. Figure 1.1 shows an overview of the major constructs of this framework and indicates how they relate to each other.

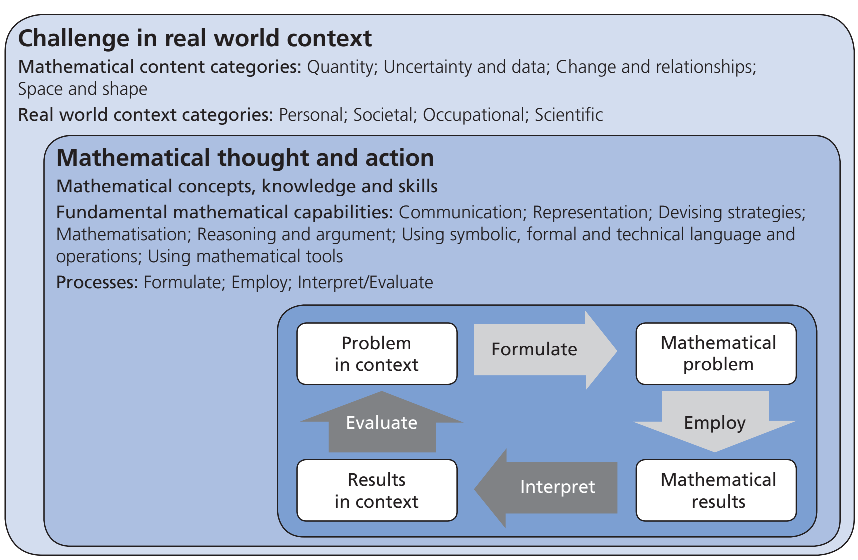
- The outer-most box in Figure 1.1 shows that mathematical literacy takes place in the context of a challenge or problem that arises in the real world. In this framework, these challenges are characterised in two ways. The context categories, which will be described in detail later in this document, identify the areas of life from which the problem arises. The context may be of a *personal* nature, involving problems or challenges that might confront an individual or one's family or peer group. The problem might instead be set in a *societal* context (focusing on one's community – whether it be local, national, or global), an *occupational* context (centred on the world of work), or a *scientific* context (relating

to the application of mathematics to the natural and technological world). A problem is also characterised by the nature of the mathematical phenomenon that underlies the challenge. The four mathematical content categories identify broad classes of phenomena that mathematics has been created to analyse. These mathematical content categories (*quantity, uncertainty and data, change and relationships,* and *space and shape*) are also identified in the outer-most box of Figure 1.1.

- To solve such contextualised problems, individuals must apply mathematical thought and action to the challenge, and the framework characterises this in three different ways. First, Figure 1.1 acknowledges the need of the individual to draw upon a variety of mathematical concepts, knowledge and skills during the work. This mathematical knowledge is drawn upon as the individual represents and communicates mathematics, devises strategies, reasons and makes arguments, and so forth. These mathematical actions are characterised in the framework in terms of seven fundamental mathematical capabilities which are listed in Figure 1.1 and described in detail later in the document. As an individual works on the problem – which may require problem formulation, employing mathematical concepts or procedures, or interpretation of a mathematical solution – the fundamental mathematical capabilities are activated successively and simultaneously, drawing on mathematical content from appropriate topics, to create a solution.
- The visual depiction of the mathematical modelling cycle in the inner-most box of Figure 1.1 portrays an idealised and simplified version of the stages through which a problem solver moves when exhibiting mathematical literacy. It shows an idealised series of stages that begin with the “problem in context”. The problem solver tries to identify the relevant mathematics in the problem situation and *formulates* the situation mathematically according to the concepts and relationships identified and simplifying assumptions made. The problem solver thus transforms the “problem in context” into a “mathematical problem” amenable to mathematical treatment. The downward-pointing arrow in Figure 1.1 depicts the work undertaken as the problem solver *employs* mathematical concepts, procedures, facts, and tools to obtain “mathematical results”. This stage typically involves mathematical reasoning, manipulation, transformation and computation. Next, the “mathematical results” need to be *interpreted* in terms of the original problem (“results in context”). This involves the problem solver interpreting, applying, and evaluating mathematical outcomes and their reasonableness in the context of a real-world-based problem. These processes of *formulating, employing,* and *interpreting* mathematics are key components of the mathematical modelling cycle and also key components of the definition of mathematical literacy. These three processes each draw on fundamental mathematical capabilities, which in turn draw on the problem solver’s detailed mathematical knowledge about individual topics.

■ Figure 1.1 ■

A model of mathematical literacy in practice



The modelling cycle is a central aspect of the PISA conception of students as active problem solvers; however, it is often not necessary to engage in every stage of the modelling cycle, especially in the context of an assessment (Niss et al., 2007). It is often the case that significant parts of the mathematical modelling cycle have been undertaken by others, and the end user carries out some of the steps of the modelling cycle, but not all of them. For example, in some cases, mathematical representations, such as graphs or equations, are given that can be directly manipulated in order to answer some question or to draw some conclusion. For this reason, many PISA items involve only parts of the modelling cycle. In reality, the

problem solver may also sometimes oscillate between the processes, returning to revisit earlier decisions and assumptions. Each of the processes may present considerable challenges, and several iterations around the whole cycle may be required.

An explicit link to a variety of contexts for problems in PISA 2012

The reference to "a variety of contexts" in the definition of mathematical literacy is purposeful and intended as a way to link to the specific contexts that are described and exemplified more fully later in this framework. The specific contexts themselves are not so important, but the four categories selected for use here (personal, occupational, societal, and scientific) do reflect a wide range of situations in which individuals may meet mathematical opportunities. The definition also acknowledges that mathematical literacy helps individuals recognise the role that mathematics plays in the world and in helping them make the kinds of well-founded judgments and decisions required of constructive, engaged, and reflective citizens.

A visible role for mathematical tools, including technology in PISA 2012

The definition of mathematical literacy explicitly includes the use of mathematical tools. These tools are physical and digital equipment, software, and calculation devices.[1] Computer-based mathematical tools are in common use in workplaces of the 21st century, and will be increasingly more prevalent as the century progresses. The nature of work-related problems and logical reasoning has expanded with these new opportunities – creating enhanced expectations for mathematical literacy.

A computer-based assessment of mathematics is an area for innovation within the PISA 2012 survey, and is being offered as an option to participating countries. Reference to mathematical tools in the definition of mathematical literacy is, therefore, particularly appropriate. The use of calculators has been permitted in all PISA mathematics surveys to date, where consistent with the policy of the participating country. While previous PISA mathematics items have been developed to be as 'calculator neutral' as possible, for some of the paper-based items presented to students in 2012, a calculator may be of positive assistance; and for the optional computer-based survey component, mathematical tools such as an online calculator will be included as part of the computer-based test material provided for some questions. Since PISA items reflect problems that arise in personal, occupational, societal, and scientific contexts, and calculators are used in all of these settings, a calculator is of assistance in some PISA items. A computer-based assessment will provide the opportunity to include a wider range of mathematics tools – such as statistical software, geometric construction and visualisation utilities, and virtual measuring instruments – into the assessment items. This will reflect the medium that increasingly more individuals use for interfacing with their world and for solving problems, and it also will provide the opportunity to assess some aspects of mathematical literacy that are not easily assessed via traditional paper-based tests.

ORGANISING THE DOMAIN

The PISA mathematics framework defines the domain of mathematics for the PISA survey and describes an approach to the assessment of the mathematical literacy of 15-year-olds. That is, PISA assesses the extent to which 15-year-old students can handle mathematics adeptly when confronted with situations and problems – the majority of which are presented in real-world contexts.

For purposes of the assessment, the PISA 2012 definition of mathematical literacy can be analysed in terms of three interrelated aspects:

- the mathematical *processes* that describe what individuals do to connect the context of the problem with mathematics and thus solve the problem, and the capabilities that underlie those processes;
- the mathematical *content* that is targeted for use in the assessment items; and
- the *contexts* in which the assessment items are located.

The following sections elaborate these aspects. In highlighting these aspects of the domain, the PISA 2012 mathematics framework helps ensure that assessment items developed for the survey reflect a range of processes, content, and contexts, so that, considered as a whole, the set of assessment items effectively operationalises what this framework defines as mathematical literacy. Several questions, based on the PISA 2012 definition of mathematical literacy lie behind the organisation of this section of the framework. They are:

- What processes do individuals engage in when solving contextual mathematical problems, and what capabilities do we expect individuals to be able to demonstrate as their mathematical literacy grows?
- What mathematical content knowledge can we expect of individuals – and of 15-year-old students in particular?
- In what contexts is mathematical literacy able to be observed and assessed?

Mathematical processes and the underlying mathematical capabilities

Mathematical processes

The definition of mathematical literacy refers to an individual's capacity to *formulate, employ*, and *interpret* mathematics. These three words, "formulate", "employ" and "interpret", provide a useful and meaningful structure for organising the mathematical processes that describe what individuals do to connect the context of a problem with the mathematics and thus solve the problem. The PISA 2012 mathematics survey will, for the first time, report results according to these mathematical processes, and this structure will provide useful and policy-relevant categories when reporting results. The categories to be used for reporting are as follows:

- *formulating* situations mathematically;
- *employing* mathematical concepts, facts, procedures, and reasoning; and
- *interpreting*, applying and evaluating mathematical outcomes.

It is important for both policy makers and those engaged more closely in the day-to-day education of students to know how effectively students are able to engage in each of these processes. The results of the PISA survey for the formulating process indicate how effectively students are able to recognise and identify opportunities to use mathematics in problem situations and then provide the necessary mathematical structure needed to formulate that contextualised problem into a mathematical form. The results of the PISA survey for the employing process indicate how well students are able to perform computations and manipulations and apply the concepts and facts that they know to arrive at a mathematical solution to a problem formulated mathematically. The results of the PISA survey for the interpreting process indicate how effectively students are able to reflect upon mathematical solutions or conclusions, interpret them in the context of a real-world problem, and determine whether the results or conclusions are reasonable. Students' facility at applying mathematics to problems and situations is dependent on skills inherent in all three of these processes, and an understanding of their effectiveness in each category can help inform both policy-level discussions and decisions being made closer to the classroom level.

Formulating situations mathematically

The word "formulate" in the mathematical literacy definition refers to individuals being able to recognise and identify opportunities to use mathematics and then provide mathematical structure to a problem presented in some contextualised form. In the process of *formulating situations mathematically*, individuals determine where they can extract the essential mathematics to analyse, set up, and solve the problem. They translate from a real-world setting to the domain of mathematics and provide the real-world problem with mathematical structure, representations, and specificity. They reason about and make sense of constraints and assumptions in the problem. Specifically, this process of *formulating situations mathematically* includes activities such as the following:

- identifying the mathematical aspects of a problem situated in a real-world context and identifying the significant variables;
- recognising mathematical structure (including regularities, relationships, and patterns) in problems or situations;
- simplifying a situation or problem in order to make it amenable to mathematical analysis;
- identifying constraints and assumptions behind any mathematical modelling and simplifications gleaned from the context;
- representing a situation mathematically, using appropriate variables, symbols, diagrams, and standard models;
- representing a problem in a different way, including organising it according to mathematical concepts and making appropriate assumptions;
- understanding and explaining the relationships between the context-specific language of a problem and the symbolic and formal language needed to represent it mathematically;
- translating a problem into mathematical language or a representation;
- recognising aspects of a problem that correspond with known problems or mathematical concepts, facts, or procedures; and
- using technology (such as a spreadsheet or the list facility on a graphing calculator) to portray a mathematical relationship inherent in a contextualised problem.

The released PISA item *PIZZAS* (see "Illustrative PISA mathematics items" at the end of this chapter) calls most heavily on students' abilities to formulate a situation mathematically. While it is indeed the case that students are also called upon to perform calculations as they solve the problem and make sense of the results of their calculations by identifying which pizza is the better value for the money, the real cognitive challenge of this item lies in being able to formulate a

mathematical model that encapsulates the concept of value for money. The problem solver must recognise that because the pizzas have the same thickness but different diameters, the focus of the analysis can be on the area of the circular surface of the pizza. The relationship between amount of pizza and amount of money is then captured in the concept of value for money, modelled as cost per unit of area. The released PISA item *ROCK CONCERT* (see "Illustrative PISA mathematics items" at the end of this chapter) is another example of an item that relies most heavily on students' abilities to formulate a situation mathematically, as it calls on students to make sense of the contextual information provided (e.g. field size and shape, the fact that the rock concert is full, and the fact that fans are standing) and translate that information into a useful mathematical form in order to estimate the number of people attending the concert.

Employing mathematical concepts, facts, procedures and reasoning

The word "employ" in the mathematical literacy definition refers to individuals being able to apply mathematical concepts, facts, procedures, and reasoning to solve mathematically-formulated problems to obtain mathematical conclusions. In the process of *employing mathematical concepts, facts, procedures and reasoning* to solve problems, individuals perform the mathematical procedures needed to derive results and find a mathematical solution (e.g. performing arithmetic computations, solving equations, making logical deductions from mathematical assumptions, performing symbolic manipulations, extracting mathematical information from tables and graphs, representing and manipulating shapes in space, and analysing data). They work on a model of the problem situation, establish regularities, identify connections between mathematical entities, and create mathematical arguments. Specifically, this process of *employing mathematical concepts, facts, procedures, and reasoning* includes activities such as:

- devising and implementing strategies for finding mathematical solutions;
- using mathematical tools, including technology, to help find exact or approximate solutions;
- applying mathematical facts, rules, algorithms, and structures when finding solutions;
- manipulating numbers, graphical and statistical data and information, algebraic expressions and equations, and geometric representations;
- making mathematical diagrams, graphs, and constructions and extracting mathematical information from them;
- using and switching between different representations in the process of finding solutions;
- making generalisations based on the results of applying mathematical procedures to find solutions; and
- reflecting on mathematical arguments and explaining and justifying mathematical results.

The released PISA unit *WALKING* (see "Illustrative PISA mathematics items" at the end of this chapter) exemplifies items that rely most heavily on students' abilities for *employing mathematical concepts, facts, procedures, and reasoning*. Both items in this unit depend upon employing a given model – a formula – to determine either the pace length (Question 1) or walking speed (Question 2). Both questions have been expressed in terms that already have mathematical structure, and students are required to perform algebraic manipulations and calculations in order to derive solutions. Similarly, the released PISA item *CARPENTER* (see "Illustrative PISA mathematics items" at the end of this chapter) relies most heavily on students *employing mathematical concepts, facts, procedures, and reasoning*. The major cognitive challenge is to devise a strategy to find information about the total length of line segments of individually unknown lengths and to reason about the comparative lengths. Individuals also have to relate the diagrams to the gardens and the perimeters to the amount of timber available, but this process of formulating is considerably less demanding than the process of reasoning about the perimeter lengths.

Interpreting, applying and evaluating mathematical outcomes

The word "interpret" used in the mathematical literacy definition focuses on the abilities of individuals to reflect upon mathematical solutions, results, or conclusions and interpret them in the context of real-life problems. This involves translating mathematical solutions or reasoning back into the context of a problem and determining whether the results are reasonable and make sense in the context of the problem. This mathematical process category encompasses both the "interpret" and "evaluate" arrows noted in the previously defined model of mathematical literacy in practice (see Figure 1.1). Individuals engaged in this process may be called upon to construct and communicate explanations and arguments in the context of the problem, reflecting on both the modelling process and its results. Specifically, this process of *interpreting, applying, and evaluating mathematical outcomes* includes activities such as:

- interpreting a mathematical result back into the real world context;
- evaluating the reasonableness of a mathematical solution in the context of a real-world problem;
- understanding how the real world impacts the outcomes and calculations of a mathematical procedure or model in order to make contextual judgments about how the results should be adjusted or applied;

- explaining why a mathematical result or conclusion does, or does not, make sense given the context of a problem;
- understanding the extent and limits of mathematical concepts and mathematical solutions; and
- critiquing and identifying the limits of the model used to solve a problem.

The released PISA item *LITTER* (see "Illustrative PISA mathematics items" at the end of this chapter) calls most heavily on students' capacity for *interpreting, applying, and evaluating* mathematical outcomes. The focus of this item is on evaluating the effectiveness of the mathematical outcome – in this case an imagined or sketched bar graph – in portraying the data presented in the item on the decomposition time of several types of litter. The item involves reasoning about the data presented, thinking mathematically about the relationship between the data and their presentation, and evaluating the result. The problem solver must and provide a reason why a bar graph is unsuitable for displaying the provided data.

Fundamental mathematical capabilities underlying the mathematical processes

A decade of experience in developing PISA items and analysing the ways in which students respond to items has revealed that there is a set of fundamental mathematical capabilities that underpins each of these reported processes and mathematical literacy in practice. The work of Mogens Niss and his Danish colleagues (Niss, 2003; Niss and Jensen, 2002; Niss and Højgaard, 2011) identified eight capabilities – referred to as "competencies" by Niss and in the 2003 framework (OECD, 2003) – that are instrumental to mathematical behaviour. The PISA 2012 framework uses a modified formulation of this set of capabilities, which condenses the number from eight to seven based on the MEG's investigation of the operation of the competencies through previously administered PISA items (Turner et al., forthcoming). There is wide recognition of the need to identify such a set of general mathematical capabilities, to complement the role of specific mathematical content knowledge in mathematics learning. Prominent examples include the eight mathematical practices of the Common Core State Standards in the United States (2010), the four key processes (representing, analysing, interpreting and evaluating, and communicating and reflecting) of the England's Mathematics National Curriculum (Qualifications and Curriculum Authority, 2007), and the process standards in the National Council of Teachers of Mathematics (NCTM) "Principles and Standards for School Mathematics" (NCTM, 2000). These cognitive capabilities are available to or learnable by individuals in order to understand and engage with the world in a mathematical way, or to solve problems. As the level of mathematical literacy possessed by an individual increases, that individual is able to draw to an increasing degree on the fundamental mathematical capabilities (Turner and Adams, 2012). Thus, increasing activation of fundamental mathematical capabilities is associated with increasing item difficulty. This observation has been used as the basis of the descriptions of different proficiency levels of mathematical literacy reported in previous PISA surveys and discussed later in this framework, in Box 1.1.

The seven fundamental mathematical capabilities used in this framework are as follows:

- *Communication*: Mathematical literacy involves *communication*. The individual perceives the existence of some challenge and is stimulated to recognise and understand a problem situation. Reading, decoding and interpreting statements, questions, tasks or objects enables the individual to form a mental model of the situation, which is an important step in understanding, clarifying and formulating a problem. During the solution process, intermediate results may need to be summarised and presented. Later on, once a solution has been found, the problem solver may need to present the solution, and perhaps an explanation or justification, to others.
- *Mathematising*: Mathematical literacy can involve transforming a problem defined in the real world to a strictly mathematical form (which can include structuring, conceptualising, making assumptions, and/or formulating a model), or interpreting or evaluating a mathematical outcome or a mathematical model in relation to the original problem. The term "mathematising" is used to describe the fundamental mathematical activities involved.
- *Representation*: Mathematical literacy very frequently involves *representations* of mathematical objects and situations. This can entail selecting, interpreting, translating between, and using a variety of representations to capture a situation, interact with a problem, or to present one's work. The representations referred to include graphs, tables, diagrams, pictures, equations, formulae, and concrete materials.
- *Reasoning and argument*: A mathematical ability that is called on throughout the different stages and activities associated with mathematical literacy is referred to as *reasoning and argument*. This capability involves logically rooted thought processes that explore and link problem elements so as to make inferences from them, check a justification that is given, or provide a justification of statements or solutions to problems.
- *Devising strategies for solving problems*: Mathematical literacy frequently requires *devising strategies for solving problems* mathematically. This involves a set of critical control processes that guide an individual to effectively recognise, formulate and solve problems. This skill is characterised as selecting or devising a plan or strategy to use mathematics

to solve problems arising from a task or context, as well as guiding its implementation. This mathematical capability can be demanded at any of the stages of the problem-solving process.

- *Using symbolic, formal and technical language and operations*: Mathematical literacy requires *using symbolic, formal and technical language and operations*. This involves understanding, interpreting, manipulating, and making use of symbolic expressions within a mathematical context (including arithmetic expressions and operations) governed by mathematical conventions and rules. It also involves understanding and utilising formal constructs based on definitions, rules and formal systems and also using algorithms with these entities. The symbols, rules and systems used will vary according to what particular mathematical content knowledge is needed for a specific task to formulate, solve or interpret the mathematics.
- *Using mathematical tools*: The final mathematical capability that underpins mathematical literacy in practice is *using mathematical tools*. Mathematical tools encompass physical tools such as measuring instruments, as well as calculators and computer-based tools that are becoming more widely available. This ability involves knowing about and being able to make use of various tools that may assist mathematical activity, and knowing about the limitations of such tools. Mathematical tools can also have an important role in communicating results. Previously it has been possible to include the use of tools in paper-based PISA surveys in only a very minor way. The optional computer-based component of the PISA 2012 mathematics assessment will provide more opportunities for students to use mathematical tools and to include observations about the way tools are used as part of the assessment.

These capabilities are evident to varying degrees in each of the three mathematical processes to be used for reporting purposes. The ways in which these capabilities manifest themselves within the three processes are described in Figure 1.2. More detail on these capabilities, particularly as they relate to item difficulty, can be found at the end of this chapter in the Box 1.1 "Fundamental mathematical capabilities and their relationship to item difficulty". In addition, each of the illustrative examples provided in the section "Illustrative PISA mathematics items" describes how the capabilities might be activated by students solving that particular problem.

Mathematical content knowledge

An understanding of mathematical content – and the ability to apply that knowledge to the solution of meaningful contextualised problems – is important for citizens in the modern world. That is, to solve problems and interpret situations in personal, occupational, societal and scientific contexts, there is a need to draw upon certain mathematical knowledge and understandings.

Mathematical structures have been developed over time as a means to understand and interpret natural and social phenomena. In schools, the mathematics curriculum is typically organised around content strands (e.g. number, algebra and geometry) and detailed topic lists that reflect historically well-established branches of mathematics and that help in defining a structured curriculum. However, outside the mathematics classroom, a challenge or situation that arises is usually not accompanied by a set of rules and prescriptions that shows how the challenge can be met. Rather it typically requires some creative thought in seeing the possibilities of bringing mathematics to bear on the situation and in formulating it mathematically. Often a situation can be addressed in different ways drawing on different mathematical concepts, procedures, facts or tools.

Since the goal of PISA is to assess mathematical literacy, an organisational structure for mathematical content knowledge is proposed based on the mathematical phenomena that underlie broad classes of problems and which have motivated the development of specific mathematical concepts and procedures. For example, mathematical phenomena such as uncertainty and change underlie many commonly occurring situations, and mathematical strategies and tools have been developed to analyse such situations. Such an organisation for content is not new, as exemplified by two well-known publications: *On the Shoulders of Giants: New Approaches to Numeracy* (Steen, 1990) and *Mathematics: The Science of Patterns* (Devlin, 1994).

Because national mathematics curricula are typically designed to equip students with knowledge and skills that address these same underlying mathematical phenomena, the outcome is that the range of content arising from organising content this way is closely aligned with that typically found in national mathematics curricula. For guidance to item writers, this framework also lists some content topics appropriate for assessing the mathematical literacy of 15-year-old students, based on analyses of national standards from eleven countries.[2]

To organise the domain of mathematics for purposes of assessing mathematical literacy, it is important to select a structure that grows out of historical developments in mathematics, that encompasses sufficient variety and depth to reveal the essentials of mathematics, and that also represents, or includes, the conventional mathematical strands in an

■ Figure 1.2 ■

Relationship between mathematical processes and fundamental mathematical capabilities

	***Formulating* situations mathematically**	***Employing* mathematical concepts, facts, procedures and reasoning**	***Interpreting*, applying and evaluating mathematical outcomes**
Communicating	Read, decode, and make sense of statements, questions, tasks, objects, images, or animations (in computer-based assessment) in order to form a mental model of the situation	Articulate a solution, show the work involved in reaching a solution and/or summarise and present intermediate mathematical results	Construct and communicate explanations and arguments in the context of the problem
Mathematising	Identify the underlying mathematical variables and structures in the real world problem, and make assumptions so that they can be used	Use an understanding of the context to guide or expedite the mathematical solving process, e.g. working to a context-appropriate level of accuracy	Understand the extent and limits of a mathematical solution that are a consequence of the mathematical model employed
Representation	Create a mathematical representation of real-world information	Make sense of, relate and use a variety of representations when interacting with a problem	Interpret mathematical outcomes in a variety of formats in relation to a situation or use; compare or evaluate two or more representations in relation to a situation
Reasoning and argument	Explain, defend or provide a justification for the identified or devised representation of a real-world situation	Explain, defend or provide a justification for the processes and procedures used to determine a mathematical result or solution Connect pieces of information to arrive at a mathematical solution, make generalisations or create a multi-step argument	Reflect on mathematical solutions and create explanations and arguments that support, refute or qualify a mathematical solution to a contextualised problem
Devising strategies for solving problems	Select or devise a plan or strategy to mathematically reframe contextualised problems	Activate effective and sustained control mechanisms across a multi-step procedure leading to a mathematical solution, conclusion, or generalisation	Devise and implement a strategy in order to interpret, evaluate and validate a mathematical solution to a contextualised problem
Using symbolic, formal and technical language and operations	Use appropriate variables, symbols, diagrams and standard models in order to represent a real-world problem using symbolic/formal language	Understand and utilise formal constructs based on definitions, rules and formal systems as well as employing algorithms	Understand the relationship between the context of the problem and representation of the mathematical solution. Use this understanding to help interpret the solution in context and gauge the feasibility and possible limitations of the solution
Using mathematical tools	Use mathematical tools in order to recognise mathematical structures or to portray mathematical relationships	Know about and be able to make appropriate use of various tools that may assist in implementing processes and procedures for determining mathematical solutions	Use mathematical tools to ascertain the reasonableness of a mathematical solution and any limits and constraints on that solution, given the context of the problem

acceptable way. Historically, with the 17th century invention of analytic geometry and calculus, mathematics became an integrated study of number, shape, change, and relationships; analysis of such phenomena as randomness and indeterminacy became instrumental to problem solving in the 19th and 20th centuries. Thus, a set of *content categories* that reflects the range of underlying mathematical phenomena was selected for the PISA 2012 framework, consistent with the categories used for previous PISA surveys.

The following list of content categories, therefore, is used in PISA 2012 to meet the requirements of historical development, coverage of the domain of mathematics and the underlying phenomena which motivate its development, and reflection of the major strands of school curricula. These four categories characterise the range of mathematical content that is central to the discipline and illustrate the broad areas of content that guide development of test items for PISA 2012:

- *change and relationships;*
- *space and shape;*
- *quantity;* and
- *uncertainty and data.*[3]

With these four categories, the mathematical domain can be organised in a way that ensures a spread of items across the domain and focuses on important mathematical phenomena, but at the same time, avoids a too fine division that would work against a focus on rich and challenging mathematical problems based on real situations. While categorisation by content category is important for item development and selection, and for reporting of assessment results, it is important to note that some specific content topics may materialise in more than one content category. For example, a released PISA item called *PIZZAS* involves determining which of two round pizzas, with different diameters and different costs but the same thickness, is the better value (see "Illustrative PISA mathematics items" at the end of this chapter to view this item and an analysis of its attributes). This item draws on several areas of mathematics, including measurement, quantification (value for money, proportional reasoning and arithmetic calculations), and change and relationships (in terms of relationships among the variables and how relevant properties change from the smaller pizza to the larger one). This item was ultimately categorised as a *change and relationships* item since the key to the problem lies in students being able to relate the change in areas of the two pizzas (given a change in diameter) and a corresponding change of price. Clearly, a different item involving circle area might be classified as a *space and shape* item. Connections between aspects of content that span these four content categories contribute to the coherence of mathematics as a discipline and are apparent in some of the assessment items selected for the PISA 2012 assessment.

The broad mathematical content categories and the more specific content topics appropriate for 15-year-old students described later in this section reflect the level and breadth of content that is eligible for inclusion on the PISA 2012 survey. Narrative descriptions of each content category and the relevance of each to solving meaningful problems are provided first, followed by more specific definitions of the kinds of content that are appropriate for inclusion in an assessment of mathematical literacy of 15-year-old students. These specific topics reflect commonalities found in the expectations set by a range of countries and educational jurisdictions. The standards examined to identify these content topics are viewed as evidence not only of what is taught in mathematics classrooms in these countries but also as indicators of what countries view as important knowledge and skills for preparing students of this age to become constructive, engaged and reflective citizens.

Descriptions of the mathematical content knowledge that characterise each of the four categories – *change and relationships, space and shape, quantity* and *uncertainty and data* – are provided below.

Change and relationships

The natural and designed worlds display a multitude of temporary and permanent relationships among objects and circumstances, where changes occur within systems of interrelated objects or in circumstances where the elements influence one another. In many cases these changes occur over time, and in other cases changes in one object or quantity are related to changes in another. Some of these situations involve discrete change; others change continuously. Some relationships are of a permanent, or invariant, nature. Being more literate about change and relationships involves understanding fundamental types of change and recognising when they occur in order to use suitable mathematical models to describe and predict change. Mathematically this means modelling the change and the relationships with appropriate functions and equations, as well as creating, interpreting, and translating among symbolic and graphical representations of relationships.

Change and relationships is evident in such diverse settings as growth of organisms, music, the cycle of seasons, weather patterns, employment levels and economic conditions. Aspects of the traditional mathematical content of functions and algebra, including algebraic expressions, equations and inequalities, tabular and graphical representations, are central in describing, modelling, and interpreting change phenomena. For example, the released PISA unit *WALKING* (see "Illustrative PISA mathematics items" at the end of this chapter) contains two items that exemplify the *change and relationships* category since the focus is on the algebraic relationships between two variables, requiring students to activate their algebraic knowledge and skills. Students are required to employ a given formula for pace length – a formula expressed in algebraic form – to determine pace length in one item and walking speed in the other. Representations of data and relationships described using statistics also are often used to portray and interpret change and relationships, and a firm grounding in the basics of number and units is also essential to defining and interpreting *change and relationships*. Some interesting relationships arise from geometric measurement, such as the way that changes in perimeter of a family of shapes might relate to changes in area, or the relationships among lengths of the sides of triangles. The released PISA item *PIZZAS* (see "Illustrative PISA mathematics items" at the end of this chapter) exemplifies *change and relationships*.

The optional computer-based assessment of mathematics in 2012 makes it possible to present students with dynamic images, multiple representations that are dynamically linked, and the opportunity to manipulate functions. For example, change over time (e.g. growth or movement) can be directly depicted in animations and simulations, and represented by linked functions, graphs and tables of data. Finding and using mathematical models of change is enhanced when individuals can explore and describe change by working with software that can graph functions, manipulate parameters, produce tables of values, experiment with geometric relationships, organise and plot data, and calculate with formulas. The capability of spreadsheets and graphing utilities to work with formulas and plot data is especially relevant.

Space and shape

Space and shape encompasses a wide range of phenomena that are encountered everywhere in our visual and physical world: patterns, properties of objects, positions and orientations, representations of objects, decoding and encoding of visual information, navigation and dynamic interaction with real shapes as well as with representations. Geometry serves as an essential foundation for *space and shape*, but the category extends beyond traditional geometry in content, meaning and method, drawing on elements of other mathematical areas such as spatial visualisation, measurement and algebra. For instance, shapes can change, and a point can move along a locus, thus requiring function concepts. Measurement formulas are central in this area. The manipulation and interpretation of shapes in settings that call for tools ranging from dynamic geometry software to Global Positioning System (GPS) software are included in this content category.

PISA assumes that the understanding of a set of core concepts and skills is important to mathematical literacy relative to *space and shape*. Mathematical literacy in the area of *space and shape* involves a range of activities such as understanding perspective (for example in paintings), creating and reading maps, transforming shapes with and without technology, interpreting views of three-dimensional scenes from various perspectives and constructing representations of shapes. The released PISA item *CARPENTER* (see "Illustrative PISA mathematics items" at the end of this chapter) belongs to this category since it deals with another key aspect of *space and shape* – properties of shapes. In this complex multiple-choice item, students are presented with four different designs for a garden bed and asked which one(s) can be edged with 32 metres of timber. This item requires the application of geometrical knowledge and reasoning. Enough information is given to enable direct calculation of the exact perimeter for three of the designs; however, inexact information is given for one design, meaning that students need to employ qualitative geometric reasoning skills.

Computer-based assessment provides students with the opportunity to manipulate dynamic representations of shapes and explore relationships within and among geometrical objects in three dimensions, which can be virtually rotated to promote an accurate mental image. Students can work with maps where zooming and rotation are possible to build up a mental picture of a place and use such tools to assist in planning routes. They can choose and use virtual tools to make measurements (e.g., of angles and line segments) on plans, images and models, and use the data in calculations. Technology allows students to integrate knowledge of geometry with visual information to build an accurate mental model. For example, to find the volume of a cup, an individual might manipulate the image to identify that it is a truncated cone, to identify the perpendicular height and where it may be measured, and to ascertain that what might look like ellipses at the top and bottom in a two-dimensional picture are actually circles in three-dimensional space.

Quantity

The notion of *quantity* may be the most pervasive and essential mathematical aspect of engaging with, and functioning in, our world. It incorporates the quantification of attributes of objects, relationships, situations and entities in the world,

understanding various representations of those quantifications, and judging interpretations and arguments based on quantity. To engage with the quantification of the world involves understanding measurements, counts, magnitudes, units, indicators, relative size, and numerical trends and patterns. Aspects of quantitative reasoning – such as number sense, multiple representations of numbers, elegance in computation, mental calculation, estimation and assessment of reasonableness of results – are the essence of mathematical literacy relative to *quantity*.

Quantification is a primary method for describing and measuring a vast set of attributes of aspects of the world. It allows for the modelling of situations, for the examination of change and relationships, for the description and manipulation of space and shape, for organising and interpreting data, and for the measurement and assessment of uncertainty. Thus mathematical literacy in the area of *quantity* applies knowledge of number and number operations in a wide variety of settings. The released PISA item *ROCK CONCERT* (see "Illustrative PISA mathematics items" at the end of this chapter) is an item exemplifying the *quantity* category. This item asks students to estimate the total number of people attending a concert, given the dimensions of the rectangular field reserved for the concert. While this item also has some elements that relate to the *space and shape* category, its primary demand comes from postulating a reasonable area for each person and using the total area available to calculate an estimated number of people attending. Alternately, given that this item is multiple-choice, students might work backwards using the area of the field and each of the response options to calculate the corresponding space per person, determining which provides the most reasonable result. Since response options are provided in terms of thousands (e.g. 2000, 5000) this item also calls on students' numerical estimation skills.

Computer-based assessment provides students with the opportunity to take advantage of the vast computational power of modern technology. It is important to note that while technology can relieve the burden of computation from individuals and free some cognitive resources to focus on meaning and strategy when solving problems, this does not remove the need for mathematically literate individuals to have a deep understanding of mathematics. An individual without such an understanding can at best use technology for routine tasks only, which is not consistent with the PISA 2012 definition of mathematical literacy. Moreover, integration of technology into the optional computer-based assessment allows for the inclusion of items that call for levels of numeric and statistical calculation that are unmanageable in the paper-based assessment.

Uncertainty and data

In science, technology and everyday life, uncertainty is a given. Uncertainty is therefore a phenomenon at the heart of the mathematical analysis of many problem situations, and the theory of probability and statistics as well as techniques of data representation and description have been established to deal with it. The *uncertainty and data* content category includes recognising the place of variation in processes, having a sense of the quantification of that variation, acknowledging uncertainty and error in measurement, and knowing about chance. It also includes forming, interpreting and evaluating conclusions drawn in situations where uncertainty is central. The presentation and interpretation of data are key concepts in this category (Moore, 1997).

There is uncertainty in scientific predictions, poll results, weather forecasts, and economic models. There is variation in manufacturing processes, test scores and survey findings, and chance is fundamental to many recreational activities enjoyed by individuals. The traditional curricular areas of probability and statistics provide formal means of describing, modelling and interpreting a certain class of uncertainty phenomena, and for making inferences. In addition, knowledge of number and of aspects of algebra such as graphs and symbolic representation contribute to facility in engaging in problems in this content category. The released PISA item *LITTER* (see "Illustrative PISA mathematics items" at the end of this chapter) is categorised as dealing with *uncertainty and data*. This item requires students to examine data presented in a table and explain why a bar graph is not suitable for displaying these data. The focus on the interpretation and presentation of data is an important aspect of the *uncertainty and data* category.

Computer-based assessment gives students the opportunity to work with larger data sets and provides the computational power and data handling capacities they need to work with such sets. Students are given the opportunity to choose appropriate tools to manipulate, analyse and represent data, and to sample from data populations. Linked representations allow students to examine and describe such data in different ways. The capacity to generate random outcomes, including numbers, enables probabilistic situations to be explored using simulations, such as the empirical likelihood of events and properties of samples.

Content topics for guiding the assessment of mathematical literacy for 15-year-old students

To effectively understand and solve contextualised problems involving *change and relationships, space and shape, quantity* and *uncertainty and data* requires drawing upon a variety of mathematical concepts, procedures, facts, and tools at an appropriate level of depth and sophistication. As an assessment of mathematical literacy, PISA strives to assess

the levels and types of mathematics that are appropriate for 15-year-old students on a trajectory to become constructive, engaged and reflective citizens able to make well-founded judgments and decisions. It is also the case that PISA, while not designed or intended to be a curriculum-driven assessment, strives to reflect the mathematics that students have likely had the opportunity to learn by the time they are 15 years old.

With an eye toward developing an assessment that is both forward-thinking yet reflective of the mathematics that 15-year-old students have likely had the opportunity to learn, analyses were conducted of a sample of mathematics standards from eleven countries to determine both what is being taught to students in classrooms around the world and also what countries deem realistic and important preparation for students as they approach entry into the workplace or admission into a higher education institution. Based on commonalities identified in these analyses, coupled with the judgment of mathematics experts, content deemed appropriate for inclusion in the assessment of mathematical literacy of 15-year-old students on PISA 2012 is described below.

The four content categories of *change and relationships, space and shape, quantity* and *uncertainty and data* serve as the foundation for identifying this range of content, yet there is not a one-to-one mapping of content topics to these categories. For example, proportional reasoning comes into play in such varied contexts as making measurement conversions, analysing linear relationships, calculating probabilities and examining the lengths of sides in similar shapes. The following content is intended to reflect the centrality of many of these concepts to all four content categories and reinforce the coherence of mathematics as a discipline. It intends to be illustrative of the content topics included in PISA 2012, rather than an exhaustive listing:

- *Functions*: The concept of function, emphasising but not limited to linear functions, their properties, and a variety of descriptions and representations of them. Commonly used representations are verbal, symbolic, tabular and graphical.
- *Algebraic expressions*: Verbal interpretation of and manipulation with algebraic expressions, involving numbers, symbols, arithmetic operations, powers and simple roots.
- *Equations and inequalities*: Linear and related equations and inequalities, simple second-degree equations, and analytic and non-analytic solution methods.
- *Co-ordinate systems*: Representation and description of data, position and relationships.
- *Relationships within and among geometrical objects in two and three dimensions*: Static relationships such as algebraic connections among elements of figures (e.g. the Pythagorean Theorem as defining the relationship between the lengths of the sides of a right triangle), relative position, similarity and congruence, and dynamic relationships involving transformation and motion of objects, as well as correspondences between two- and three-dimensional objects.
- *Measurement*: Quantification of features of and among shapes and objects, such as angle measures, distance, length, perimeter, circumference, area and volume.
- *Numbers and units*: Concepts, representations of numbers and number systems, including properties of integer and rational numbers, relevant aspects of irrational numbers, as well as quantities and units referring to phenomena such as time, money, weight, temperature, distance, area and volume, and derived quantities and their numerical description.
- *Arithmetic operations*: The nature and properties of these operations and related notational conventions.
- *Percents, ratios and proportions*: Numerical description of relative magnitude and the application of proportions and proportional reasoning to solve problems.
- *Counting principles*: Simple combinations and permutations.
- *Estimation*: Purpose-driven approximation of quantities and numerical expressions, including significant digits and rounding.
- *Data collection, representation and interpretation*: Nature, genesis and collection of various types of data, and the different ways to represent and interpret them.
- *Data variability and its description*: Concepts such as variability, distribution and central tendency of data sets, and ways to describe and interpret these in quantitative terms.
- *Samples and sampling*: Concepts of sampling and sampling from data populations, including simple inferences based on properties of samples.
- *Chance and probability*: Notion of random events, random variation and its representation, chance and frequency of events, and basic aspects of the concept of probability.

Contexts

An important aspect of mathematical literacy is that mathematics is engaged in solving a problem set in a context. The context is the aspect of an individual's world in which the problems are placed. The choice of appropriate mathematical strategies and representations is often dependent on the context in which a problem arises. Being able to work within a context is widely appreciated to place additional demands on the problem solver (see Watson and Callingham, 2003, for findings about statistics). For the PISA survey, it is important that a wide variety of contexts are used. This offers the possibility of connecting with the broadest possible range of individual interests and with the range of situations in which individuals operate in the 21st century.

For purposes of the PISA 2012 mathematics framework, four context categories have been defined and are used to classify assessment items developed for the PISA survey:

- *Personal*: Problems classified in the personal context category focus on activities of one's self, one's family or one's peer group. The kinds of contexts that may be considered personal include (but are not limited to) those involving food preparation, shopping, games, personal health, personal transportation, sports, travel, personal scheduling and personal finance. The released PISA item *PIZZAS* (see "Illustrative PISA mathematics items" at the end of this chapter) is set in a personal context since the question posed by the item is which pizza provides the purchaser with the better value for the money. Similarly, the released PISA unit *WALKING* (see "Illustrative PISA mathematics items" at the end of this chapter) contains two items that reflect a personal context. The first item involves applying a mathematical formula to determine of an individual's pace length, while the second item involves the application of the same formula to determine of another individual's walking speed.
- *Occupational*: Problems classified in the occupational context category are centred on the world of work. Items categorised as occupational may involve (but are not limited to) such things as measuring, costing and ordering materials for building, payroll/accounting, quality control, scheduling/inventory, design/architecture and job-related decision making. Occupational contexts may relate to any level of the workforce, from unskilled work to the highest levels of professional work, although items in the PISA survey must be accessible to 15-year-old students. The released PISA item *CARPENTER* (see "Illustrative PISA mathematics items" at the end of this chapter) is categorised as occupational as it deals with a work task of a carpenter to construct a border around a garden bed. An item requiring similar mathematical analysis to the *PIZZAS* item discussed earlier, which presented the situation from the point of view of the pizza seller instead of the purchaser, would be placed in the occupational category.
- *Societal*: Problems classified in the societal context category focus on one's community (whether local, national or global). They may involve (but are not limited to) such things as voting systems, public transport, government, public policies, demographics, advertising, national statistics and economics. Although individuals are involved in all of these things in a personal way, in the societal context category the focus of problems is on the community perspective. The released PISA item *ROCK CONCERT* (see "Illustrative PISA mathematics items" at the end of this chapter) is an example of an item categorised as societal since it is set at the level of the rock concert organisation, even though it draws on the personal experience of being in crowds.
- *Scientific*: Problems classified in the scientific category relate to the application of mathematics to the natural world and issues and topics related to science and technology. Particular contexts might include (but are not limited to) such areas as weather or climate, ecology, medicine, space science, genetics, measurement and the world of mathematics itself. The released PISA item *LITTER* (see "Illustrative PISA mathematics items" at the end of this chapter) is an example of an item set in a scientific context, since its focus is related to scientific issues pertaining to the environment, and specifically to data on decomposition time. Items that are intramathematical, where all the elements involved belong in the world of mathematics, fall within the scientific context.

PISA assessment items are arranged in units that share stimulus material. It is therefore usually the case that all items in the same unit belong to the same context category. Exceptions do arise; for example stimulus material may be examined from a personal point of view in one item and a societal point of view in another. When an item involves only mathematical constructs without reference to the contextual elements of the unit within which it is located, it is allocated to the context category of the unit. In the unusual case of a unit involving only mathematical constructs and being without reference to any context outside of mathematics, the unit is assigned to the scientific context category.

Using these context categories provides the basis for selecting a mix of item contexts and ensures that the assessment reflects a broad range of uses of mathematics, ranging from everyday personal uses to the scientific demands of global problems. Moreover it is important that each context category be populated with assessment items having a broad range of item difficulties. Given that the major purpose of these context categories is to challenge students in a broad range of

problem contexts, each category should contribute substantially to the measurement of mathematical literacy. It should not be the case that the difficulty level of assessment items representing one context category is systematically higher or lower than the difficulty level of assessment items in another category.

In identifying contexts that may be relevant, it is critical to keep in mind that a purpose of the assessment is to gauge the use of mathematical content knowledge, processes, and capabilities that students have acquired by age 15. Contexts for assessment items, therefore, are selected in light of relevance to students' interests and lives, and the demands that will be placed upon them as they enter society as constructive, engaged and reflective citizens. National project managers from countries participating in the PISA survey are involved in judging the degree of such relevance.

ASSESSING MATHEMATICAL LITERACY

In this section, the approach taken to implement the elements of the framework described in previous sections into the PISA survey for 2012, is outlined. This includes the structure of the mathematics component of the PISA survey, the reporting of levels of mathematical proficiency, the attitudes to be investigated that relate to mathematical proficiency, and arrangements for the optional computer-based survey component for mathematics.

Structure of the PISA 2012 mathematics assessment

In accordance with the definition of mathematical literacy, assessment items used in any instruments that are developed as part of the PISA survey, both paper-based and computer-based, are set within a context. Items involve the application of important mathematical concepts, knowledge, understandings and skills (mathematical content knowledge) at the appropriate level for 15-year-old students, as described earlier. The framework is used to guide the structure and content of the assessment, and it is important that the survey instruments, both paper-based and computer-based, include an appropriate balance of items reflecting the components of the mathematical literacy framework.

Desired distribution of score points by mathematical process

In addition, assessment items in the PISA 2012 mathematics survey can be assigned to one of three mathematical processes. The goal in constructing the assessment is to achieve a balance that provides approximately equal weighting between the two processes that involve making a connection between the real world and the mathematical world and the process that calls for students to be able to work on a mathematically formulated problem.

Table 1.1

Approximate distribution of score points in mathematics, by process category

Process category	Percentage of score points
Formulating situations mathematically	Approximately 25
Employing mathematical concepts, facts, procedures and reasoning	Approximately 50
Interpreting, applying and evaluating mathematical outcomes	Approximately 25
TOTAL	100

It is important to note that items in each process category should have a range of difficulty and mathematical demand.

Desired distribution of score points by content category

PISA mathematics items are selected to reflect the mathematical content knowledge described earlier in this framework. The items selected for PISA 2012 are distributed across the four content categories, as shown in Table 1.2. The goal in constructing the survey is a distribution of items with respect to content category that provides as balanced a distribution of score points as possible, since all of these domains are important for constructive, engaged and reflective citizens.

Table 1.2

Approximate distribution of score points in mathematics, by content category

Content category	Percentage of score points
Change and relationships	Approximately 25
Space and shape	Approximately 25
Quantity	Approximately 25
Uncertainty and data	Approximately 25
TOTAL	100

It is important to note that items in each content category should have a range of difficulty and mathematical demand.

Desired distribution of score points by context category

For PISA 2012, each item is set in one of four context categories. The items selected for the PISA 2012 mathematics survey represent a spread across these context categories, as described in Table 1.3. With this balanced distribution, no single context type is allowed to dominate, providing students with items that span a broad range of individual interests and a range of situations that they might expect to encounter in their lives.

Table 1.3
Approximate distribution of score points in mathematics, by context category

Context category	Percentage of score points
Personal	Approximately 25
Occupational	Approximately 25
Societal	Approximately 25
Scientific	Approximately 25
TOTAL	100

It is important to note that items in each context category should have a range of difficulty and mathematical demand.

A range of item difficulties

The PISA 2012 mathematics survey includes items with a wide range of difficulties, paralleling the range of abilities of 15-year-old students. It includes items that are challenging for the most able students, and items that are suitable for the least able students assessed in mathematics. From a psychometric perspective, a survey that is designed to measure a particular cohort of individuals is most effective and efficient when the difficulty of assessment items matches the ability of the measured subjects. Furthermore, the described proficiency scales that are used as a central part of the reporting of PISA outcomes can only include useful details for all students if the items from which the proficiency descriptions are drawn span the range of abilities described. The proficiency scales are based on increasing levels of activation of the fundamental mathematical capabilities, described fully in the Box 1.1 "Fundamental mathematical capabilities and their relationship to item difficulty". Previous PISA cycles have shown that collectively these capabilities are indicators of cognitive demand, and thus contribute centrally to item difficulty (Turner, 2012; Turner et al., forthcoming). A scale for PISA 2012 was developed after the field test and based on a description of the required activation of these capabilities. This scale provides an empirical measure of the cognitive demand for each item.

Structure of the survey instrument

The paper-based instruments for the PISA 2012 survey contain a total of 270 minutes of mathematics material. The material is arranged in nine clusters of items, with each cluster representing 30 minutes of testing time. Of this total, three clusters (representing 90 minutes of test time) comprise link material used in previous PISA surveys, four "standard" clusters (representing 120 minutes of test time) comprise new material having a wide range of difficulty, and two "easy" clusters (representing 60 minutes of test time) are devoted to material with a lower level of difficulty.

Each participating country uses seven of the clusters: the three clusters of link material, two of the new 'standard' clusters, and either the other two "standard" clusters or the two "easy" clusters. The provision of "easy" and "standard" clusters allows for better targeting of the assessment for each of the participating countries; however, the items are scaled in such a way that a country's score will not be affected if it chooses to administer either the "easy" or additional "standard" clusters. The item clusters are placed in test booklets according to a rotated test design, with each form containing four clusters of material from the mathematics, reading and science domains. Each student does one form, representing a total testing time of 120 minutes.

The optional computer-based component (CBAM) contains a total of 80 minutes of mathematics material. The material is arranged in four clusters of items, with each cluster representing 20 minutes of testing time. This material is arranged in a number of rotated test forms along with other material for computer delivery, with each form containing two clusters. Each student does one form, representing a total testing time of 40 minutes.

Design of the PISA 2012 mathematics items

Three item format types are used in the paper-based component to assess mathematical literacy in PISA 2012: open constructed-response, closed constructed-response and selected-response (multiple-choice) items. Open constructed-response items require a somewhat extended written response from a student. Such items also may ask the student to show the steps taken or to explain how the answer was reached. These items require trained experts to manually code student responses. Closed constructed-response items provide a more structured setting for presenting problem solutions, and they produce a student response that can be easily judged to be either correct or incorrect. Often student responses to questions of this type can be keyed into data capture software, and coded automatically, but some must be manually coded by trained experts. The most frequent closed constructed-responses are single numbers. Selected-response items require the choice of one or more responses from a number of response options. Responses to these questions can usually be automatically processed. About equal numbers of each of these item format types are being used to construct the survey instruments.

In the optional computer-based component, additional item format types are possible. A computer-based environment lends itself to a wider range of response modes than does paper, as well as facilitating assessment of some aspects of mathematical literacy, such as the manipulation and rotation of representations of three-dimensional shapes, which cannot so readily be assessed on paper. A computer-based assessment enables item presentation to be enhanced. For example there may be a moving stimulus, representations of three-dimensional objects that can be rotated, or more flexible access to relevant information or data. Item formats that permit a wider range of response types are also possible. For example, drag-and-drop items or the use of hot spots on an image may allow students to respond to more items non-verbally, giving a more rounded picture of mathematical literacy that is less bound to language. Some interactivity may be possible. Additionally, the possibility of automated response coding may replace some manual work; more importantly, it may facilitate coding of features in student-constructed drawings, displays and procedures that are currently impractical to code (Stacey and Wiliam, forthcoming).

The PISA mathematics survey is composed of assessment *units* comprising verbal stimulus material and often other information such as tables, charts, graphs or diagrams, plus one or more items that are linked to this common stimulus material. This format gives students the opportunity to become involved with a context or problem by responding to a series of related items. However, the measurement model used to analyse PISA data assumes item independence, so whenever units comprising more than one item are used, the objective of item writers is to ensure maximum possible independence among the items. PISA employs this unit structure to facilitate the employment of contexts that are as realistic as possible, and that reflect the complexity of real situations, while making efficient use of testing time. However, it is important to ensure that there is an adequate range of contexts so that bias due to choice of contexts is minimised and item independence is maximised. A balance between these two competing demands is therefore sought in developing the PISA survey instruments.

Items selected for inclusion in the PISA survey represent a broad range of difficulties, to match the wide ability range of students participating in the assessment. In addition, all the major categories of the assessment (the content categories, the process categories, and the context categories) are represented, to the degree possible, with items of a wide range of difficulties. Item difficulties are established as one of a number of measurement properties in an extensive field trial prior to item selection for the main PISA survey. Items are selected for inclusion in the PISA survey instruments based on their fit with framework categories and their measurement properties.

In addition, the level of reading required to successfully engage with an item is considered very carefully in item development and selection. A goal in item development is to make the wording of items as simple and direct as possible. Care is also taken to avoid item contexts that would create a cultural bias, and all choices are checked with national teams. Translation of the items into many languages is conducted very carefully, with extensive back-translation and other protocols. Attention to item bias is even more critical in PISA 2012 since the inclusion of the optional computer-based component may present new challenges to students who have not had access to computers in their mathematics classrooms.

Mathematical tools

PISA policy allows students to use calculators in the paper-based components as they are normally used in school. This represents the most authentic assessment of what students can achieve, and provides the most informative comparison of the performance of education systems. A system's choice to allow students to access and use calculators is no different, in principle, from other instructional policy decisions that are made by systems that are not controlled by PISA. In 2012, for the first time in a PISA mathematics assessment, some of the items written for paper-based delivery will be constructed in such a way that a calculator will likely make the calculations required quicker and easier – meaning that for some assessment items, it is likely that availability of a calculator will be an advantage for many students. In the paper-based component of PISA 2012, functionalities beyond the arithmetic functionality of a basic calculator will not be required.

In the optional computer-based component of PISA 2012, students will be given access to an online calculator and/or software with equivalent functionality for items where this could be relevant. Students can also have access to a

hand-held calculator, as approved for use by 15-year-old students in their respective school systems. Other tools may also be provided as part of the test delivery system, such as virtual measuring devices, some basic spreadsheet functionality, and various graphic presentation and visualisation tools.

Item scoring

Although the majority of the items are dichotomously scored (that is, responses are awarded either credit or no credit), the open constructed-response items can sometimes involve partial credit scoring, which allows responses to be assigned credit according to differing degrees of "correctness" of responses. For each such item, a detailed coding guide that allows for full credit, partial credit or no credit is provided to persons trained in the coding of student responses across the range of participating countries to ensure coding of responses is done in a consistent and reliable way.

Reporting proficiency in mathematics

The outcomes of the PISA mathematics survey are reported in a number of ways. Estimates of overall mathematical proficiency are obtained for sampled students in each participating country, and a number of proficiency levels are defined. Descriptions of the degree of mathematical literacy typical of students in each level are also developed. In addition, aspects of overall mathematical proficiency are identified that will be of policy relevance to participating countries, separate estimates are obtained for students in relation to those aspects, and proficiency descriptions are also developed for the different levels defined on those scales. Aspects of potential use for reporting purposes can be defined in a variety of ways. For PISA 2003, scales based on the four broad content categories were developed. In Figure 1.3, descriptions for the six proficiency levels reported for the overall PISA mathematics scale in 2003, 2006 and 2009 are presented. These form the basis for the PISA 2012 mathematics scale.

■ Figure 1.3 ■

Proficiency scale descriptions for mathematics (2003-2009)

Level	
6	At Level 6 students can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply their insight and understandings along with a mastery of symbolic and formal mathematical operations and relationships to develop new approaches and strategies for attacking novel situations. Students at this level can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments and the appropriateness of these to the original situations.
5	At Level 5 students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations and insight pertaining to these situations. They can reflect on their actions and formulate and communicate their interpretations and reasoning.
4	At Level 4 students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Students at this level can utilise well-developed skills and reason flexibly, with some insight, in these contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments and actions.
3	At Level 3 students can execute clearly described procedures, including those that require sequential decisions. They can select and apply simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They can develop short communications when reporting their interpretations, results and reasoning.
2	At Level 2 students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results.
1	At Level 1 students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli.

As well as the overall mathematics scale, three additional described proficiency scales are developed after the field trial and are then reported, based on the three mathematical processes described earlier – *formulating situations*

mathematically; employing mathematical concepts, facts, procedures, and reasoning; and *interpreting, applying and evaluating mathematical outcomes*.

Fundamental mathematical capabilities play a central role in defining what it means to be at different levels of the scales for mathematical literacy overall and for each of the reported processes – they define growing proficiency for all these aspects of mathematical literacy. For example, in the proficiency scale description for Level 4 (see Figure 1.3), the second sentence highlights aspects of mathematising and representation that are evident at this level. The final sentence highlights the characteristic communication, reasoning and argument of Level 4, providing a contrast with the short communications and lack of argument of Level 3 and the additional reflection of Level 5. The Box 1.1 "Fundamental mathematical capabilities and their relationship to item difficulty" at the end of this chapter describes the fundamental mathematical capabilities and the relationship each one has to development across levels of mathematical proficiency. In an earlier section of this framework and in Figure 1.2, each of the mathematical processes was described in terms of the fundamental mathematical capabilities that individuals might activate when engaging in that process.

For continuity with the reporting of outcomes of the 2003 survey when mathematics was last the major PISA survey domain and because of its usefulness for providing information for policy decisions, scales will also be reported based on the four content categories: *quantity, space and shape, change and relationships,* and *uncertainty and data*. These scales will continue to be interesting for countries since they can show profiles in aspects of mathematical proficiency resulting from specific curricular emphases.

Attitudes towards mathematics

Individuals' attitudes, beliefs and emotions play a significant role in their interest and response to mathematics in general, and their employment of mathematics in their individual lives. Students who feel more confident with mathematics, for example, are more likely than others to use mathematics in the various contexts that they encounter. Students who have positive emotions towards mathematics are in a position to learn mathematics better than students who feel anxiety towards that subject. Therefore, one goal of mathematics education is for students to develop attitudes, beliefs and emotions that make them more likely to successfully use the mathematics they know, and to learn more mathematics, for personal and social benefit.

The attention the PISA 2012 assessment of mathematics gives to these variables is based on claims that the development of positive attitudes, emotions and beliefs towards mathematics is in itself a valuable outcome of schooling and predisposes students to use mathematics in their lives; and also that such variables may contribute to explaining differences in the achievement of mathematical literacy. The PISA survey therefore includes items related to these variables. In addition, the PISA survey measures a range of background variables that enable the reporting and analysis of mathematical literacy for important subgroups of students (e.g. by gender, language or migration status).

To gather background information, students and the principals of their schools are asked to respond to background questionnaires of around 20 to 30 minutes in length. These questionnaires are central to the analysis and reporting of results in terms of a range of student and school characteristics.

Two broad areas of students' attitudes towards mathematics that dispose them to productive engagement in mathematics are identified as being of potential interest as an adjunct to the PISA 2012 mathematics assessment. These are students' interest in mathematics and their willingness to engage in it.

Interest in mathematics has components related to present and future activity. Relevant questions focus on students' interest in mathematics at school, whether they see it as useful in real life as well as their intentions to undertake further study in mathematics and to participate in mathematics-oriented careers. There is international concern about this area, because in many participating countries there is a decline in the percentage of students who are choosing mathematics-related future studies, whereas at the same time there is a growing need for graduates from these areas.

Students' willingness to do mathematics is concerned with the attitudes, emotions and self-related beliefs that dispose students to benefit, or prevent them from benefitting, from the mathematical literacy that they have achieved. Students who enjoy mathematical activity and feel confident to undertake it are more likely to use mathematics to think about the situations that they encounter in the various facets of their lives, inside and outside school. The constructs from the PISA survey that are relevant to this area include the emotions of enjoyment, confidence and (lack of) mathematics anxiety, and the self-related beliefs of self-concept and self-efficacy. A recent analysis of the subsequent progress of young Australians who scored poorly on PISA at age 15 found that those who "recognise the value of mathematics for their

future success are more likely to achieve this success, and that includes being happy with many aspects of their personal lives as well as their futures and careers" (Thomson and Hillman, 2010, p. 31). The study recommends that a focus on the practical applications of mathematics in everyday life may help improve the outlook for these low-achieving students.

The student questionnaire also includes sets of items related to *opportunity to learn*. There are items concerning student experience with applied mathematics problems of various types, student familiarity with mathematical concepts by name (including measures to guard against overclaiming) and student experience in class or tests with PISA style items. These measures will allow deeper analysis of the PISA results.

The results of the PISA 2012 survey will provide important information for educational policy makers in the participating countries about both the achievement-related and attitude-related outcomes of schooling. By combining information from the PISA assessment of mathematical literacy and the survey of attitudes, emotions and beliefs that predispose students to use their mathematical literacy, a more complete picture emerges.

Optional computer-based assessment of mathematics

PISA 2012 includes a computer-based assessment of mathematics.[4] While the computer-based assessment is optional for participating countries (given countries' varied technological capacities), there are two aspects to the rationale for including a computer-based mathematics assessment in PISA 2012. First, computers are now so commonly used in the workplace and in everyday life that a level of competency in mathematical literacy in the 21st century includes usage of computers (Hoyles et al., 2002). Computers now touch the lives of individuals around the world as they engage in their personal, societal, occupational and scientific endeavours. They offer tools for – among other things – computation, representation, visualisation, modification, exploration and experimentation on, of and with a large variety of mathematical objects, phenomena and processes. The definition of mathematical literacy for PISA 2012 recognises the important role of computer-based tools by noting that mathematically literate individuals are expected to use these in their endeavours to describe, explain, and predict phenomena. In this definition, the word "tool" refers to calculators and computers, as well as to other physical objects such as rulers and protractors used for measuring and construction. A second consideration is that the computer provides a range of opportunities for designers to write test items that are more interactive, authentic and engaging (Stacey and Wiliam, forthcoming). These opportunities include the ability to design new item formats (e.g. drag-and-drop), to present students with real-world data (such as a large, sortable dataset), or to use colour and graphics to make the assessment more engaging.

In response to these phenomena, an optional computer-based assessment of mathematics is a major area for innovation in the PISA 2012 assessment. Specially designed PISA units are presented on a computer, and students respond on the computer. They are also able to use pencil and paper to assist their thinking processes. Future PISA cycles may feature more sophisticated computer-based items, as developers and item writers become more fully immersed in computer-based assessment. Indeed, PISA 2012 represents only a starting point for the possibilities of the computer-based assessment of mathematics.

Making use of enhancements offered by computer technology results in assessment items that are more engaging to students, more colourful, and easier to understand. For example, students may be presented with a moving stimulus, representations of three-dimensional objects that can be rotated, or more flexible access to relevant information. New item formats, such as those calling for students to 'drag and drop' information or use 'hot spots' on an image, are designed to engage students, permit a wider range of response types and give a more rounded picture of mathematical literacy.

Investigations show that the mathematical demands of work increasingly occur in the presence of electronic technology so that mathematical literacy and computer use are melded together (Hoyles et al., 2002). For employees at all levels of the workplace, there is now an interdependency between mathematical literacy and the use of computer technology, and the computer-based component of the PISA survey provides opportunities to explore this relationship. A key challenge is to distinguish the mathematical demands of a PISA computer-based item from demands unrelated to mathematical proficiency, such as the information and communications technology (ICT) demands of the item, and new presentation formats. On the optional computer-based PISA 2012 survey it is important that the focus is on ensuring that the demand associated with the use of a tool in a particular item is significantly lower than the demand associated with the mathematics. Research has been conducted on the impact a computer-based testing environment has on students' performance (Bennett, 2003; Bennett et al., 2008; Mason et al., 2001; Richardson et al., 2002; Sandene et al., 2008), and the PISA 2012 survey provides an opportunity to further this knowledge, particularly to inform development of future computer-based tests for 2015 and beyond. By design, not all computer-based items will use new item formats, which might be helpful in monitoring the (positive or negative) impact that new item formats have on performance.

In order to establish control over the range of computer-based features of the test, for each item three aspects are described:

The mathematical competencies being tested: These comprise aspects of mathematical literacy applicable in any environment, not just computer environments, and are being tested in every computer-based assessment item.

Competencies that cover aspects of mathematics and ICT: These require knowledge of doing mathematics with the assistance of a computer or handheld device. These are being tested in some – but not all – computer-based assessment items. The computer-based test may include assessments of the following competencies:

- making a chart from data, including from a table of values (e.g. pie chart, bar chart, line graph) using simple 'wizards';
- producing graphs of functions and using the graphs to answer questions about the functions;
- sorting information and planning efficient sorting strategies;
- using hand-held or on-screen calculators;
- using virtual instruments such as an on-screen ruler or protractor; and
- transforming images using a dialog box or mouse to rotate, reflect or translate the image.

ICT skills: Just as pencil and paper assessments rely on a set of fundamental skills for working with printed materials, computer-based assessments rely on a set of fundamental skills for using computers. These include knowledge of basic hardware (e.g. keyboard and mouse) and basic conventions (e.g. arrows to move forward and specific buttons to press to execute commands). The intention is to keep such skills to a minimal core level in every computer-based assessment item.

SUMMARY

The aim of PISA with regard to mathematical literacy is to develop indicators that show how effectively countries are preparing students to use mathematics in every aspect of their personal, civic and professional lives, as part of their constructive, engaged and reflective citizenship. To achieve this, PISA has developed a definition of mathematical literacy and an assessment framework that reflects the important components of this definition. The mathematics assessment items developed and selected for inclusion in PISA 2012, based on this definition and framework, are intended to reflect a balance of relevant mathematical processes, mathematical content and contexts. These items are intended to determine how students can use what they have learnt. They call for students to use the content they know by engaging in processes and applying the capabilities they possess to solve problems that arise out of real-world experiences. The assessment provides problems in a variety of item formats with varying degrees of built-in guidance and structure, but the emphasis is on authentic problems where students must do the thinking themselves.

Box 1.1 **Fundamental mathematical capabilities and their relationship to item difficulty**

A good guide to the empirical difficulty of items can be obtained by considering which aspects of the fundamental mathematical capabilities are required for planning and executing a solution (Turner, 2012; Turner and Adams, 2012; Turner et al., forthcoming). The easiest items will require the activation of few capabilities and in a relatively straightforward way. The hardest items require complex activation of several capabilities. Predicting difficulty requires consideration of both the number of capabilities and the complexity of activation required. The sections below describe characteristics which make the activation of a single capability more or less complex (see also Turner, 2012).

Communication: Various factors determine the level and extent of the communication demand of a task, and the capability of an individual to meet these demands indicates the extent to which they possess the communication capability. For the receptive aspects of communication, these factors include the length and complexity of the text or other object to be read and interpreted, the familiarity of the ideas or information referred to in the text or object, the extent to which the information required needs to be disentangled from other information, the ordering of information and whether this matches the ordering of the thought processes required to interpret and use the information, and the extent to which there are different elements (such as text, graphic elements, graphs, tables, charts) that need to be interpreted in relation to each other. For the expressive aspects of communication, the lowest level of complexity is observed in tasks that demand simply provision of a numeric answer. As the requirement for a more extensive expression of a solution is added, for example when a verbal or written explanation or justification of the result is required, the communication demand increases.

Mathematising: In some tasks, mathematisation is not required – either the problem is already in a sufficiently mathematical form, or the relationship between the model and the situation it represents is not needed to solve the

...

problem. The demand for mathematisation arises in its least complex form when the problem solver needs to interpret and infer directly from a given model, or to translate directly from a situation into mathematics (e.g. to structure and conceptualise the situation in a relevant way, to identify and select relevant variables, collect relevant measurements, and/or make diagrams). The mathematisation demand increases with additional requirements to modify or use a given model to capture changed conditions or interpret inferred relationships; to choose a familiar model within limited and clearly articulated constraints; or to create a model where the required variables, relationships and constraints are explicit and clear. At an even higher level, the mathematisation demand is associated with the need to create or interpret a model in a situation where many assumptions, variables, relationships and constraints are to be identified or defined, and to check that the model satisfies the requirements of the task; or, to evaluate or compare models.

Representation: This mathematical capability is called on at the lowest level with the need to directly handle a given familiar representation, for example going directly from text to numbers, or reading a value directly from a graph or table. More cognitively demanding representation tasks call for the selection and interpretation of one standard or familiar representation in relation to a situation, and at a higher level of demand still when they require translating between or using two or more different representations together in relation to a situation, including modifying a representation; or when the demand is to devise a straightforward representation of a situation. Higher level cognitive demand is marked by the need to understand and use a non-standard representation that requires substantial decoding and interpretation; to devise a representation that captures the key aspects of a complex situation; or to compare or evaluate different representations.

Reasoning and argument: In tasks of very low demand for activation of this capability, the reasoning required may involve simply following the instructions given. At a slightly higher level of demand, items require some reflection to connect different pieces of information in order to make inferences (e.g. to link separate components present in the problem, or to use direct reasoning within one aspect of the problem). At a higher level, tasks call for the analysis of information in order to follow or create a multi-step argument or to connect several variables; or to reason from linked information sources. At an even higher level of demand, there is a need to synthesise and evaluate information, to use or create chains of reasoning to justify inferences, or to make generalisations drawing on and combining multiple elements of information in a sustained and directed way.

Devising strategies: In tasks with a relatively low demand for this capability, it is often sufficient to take direct actions, where the strategy needed is stated or obvious. At a slightly higher level of demand, there may be a need to decide on a suitable strategy that uses the relevant given information to reach a conclusion. Cognitive demand is further heightened with the need to devise and construct a strategy to transform given information to reach a conclusion. Even more demanding tasks call for the construction of an elaborated strategy to find an exhaustive solution or a generalised conclusion; or to evaluate or compare different possible strategies.

Using symbolic, formal and technical language and operations: The demand for activation of this capability varies enormously across tasks. In the simplest tasks, no mathematical rules or symbolic expressions need to be activated beyond fundamental arithmetic calculations, operating with small or easily tractable numbers. Work with more demanding tasks may involve sequential arithmetic calculations or direct use of a simple functional relationship, either implicit or explicit (e.g. familiar linear relationships); use of formal mathematical symbols (e.g. by direct substitution or sustained arithmetic calculations involving fractions and decimals); or an activation and direct use of a formal mathematical definition, convention or symbolic concept. Further increased cognitive demand is characterised by the need for explicit use and manipulation of symbols (e.g. by algebraically rearranging a formula), or by activation and use of mathematical rules, definitions, conventions, procedures or formulas using a combination of multiple relationships or symbolic concepts. A yet higher level of demand is characterised by the need for a multi-step application of formal mathematical procedures, working flexibly with functional or involved algebraic relationships, or using both mathematical technique and knowledge to produce results.

Using mathematical tools: Tasks and activities involving a relatively low level of demand for this capability may require direct use of familiar tools, such as a measuring instrument, in situations where use of those tools is well-practised. Higher levels of demand arise when using the tool involves a sequence of processes, or linking different information using the tool, and when familiarity of the tools themselves is lower or when the situation in which the application of the tool is required is less familiar. Further increased demand is seen when the tool is to be used to process and relate multiple data elements, when the application of a tool is needed in a situation quite different from familiar applications, when the tool itself is complex with multiple affordances, and when there is a need for reflection to understand and evaluate the merits and limitations of the tool.

ILLUSTRATIVE PISA MATHEMATICS ITEMS

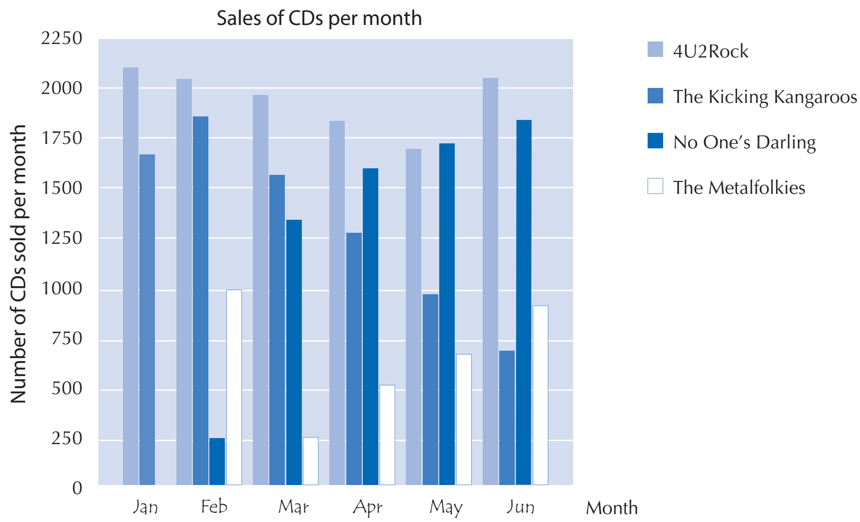
The following released PISA items are intended to illustrate relevant aspects and nuances of the PISA 2012 framework. The seven items were selected to represent a spread across item type, process, content and context, as well as to describe the activation of the fundamental mathematical capabilities, but they are not intended to represent the full range of any particular aspect.

CHARTS

The first illustrative unit is titled *CHARTS*. It comprises stimulus information in the form of text and a bar graph that represents music CD sales for four bands over a period of six months, and three simple multiple choice items (Figure 1.4).

■ Figure 1.4 ■
Items for the unit *CHARTS*

In January, the new CDs of the bands *4U2Rock* and *The Kicking Kangaroos* were released. In February, the CDs of the bands *No One's Darling* and *The Metalfolkies* followed. The following graph shows the sales of the bands' CDs from January to June.



QUESTION 1

How many CDs did the band *The Metalfolkies* sell in April?

A. 250

B. 500

C. 1 000

D. 1 270

QUESTION 2

In which month did the band *No One's Darling* sell more CDs than the band *The Kicking Kangaroos* for the first time?

A. No month

B. March

C. April

D. May

QUESTION 3

The manager of The *Kicking Kangaroos* is worried because the number of their CDs that sold decreased from February to June.

What is the estimate of their sales volume for July if the same negative trend continues?

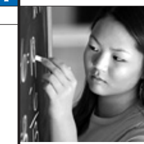
A. 70 CDs

B. 370 CDs

C. 670 CDs

D. 1 340 CDs

In preparing national versions, PISA countries were expected to replace the band names with fictitious names suitable for their local context.



CHARTS was used in the PISA 2012 main survey. The three items of *CHARTS* each lie in the *uncertainty and data* content category, since they ask students to read, interpret and use data presented in a mathematical graphical form. They each lie in the *societal* context category, since the data relate to public information about music sales, the kind that might be found in a newspaper, music magazine or on line. The first two questions are examples of the *interpreting, applying and evaluating mathematical outcomes* process category, since these questions involve interpreting the mathematical information presented in the chart in relation to context features represented; while the third question fits the *employing mathematical facts, concepts, procedures and reasoning* category because its focus is on applying procedural knowledge to manipulate the mathematical representation in order to make a further inference. The three questions were among the easiest questions used in the PISA 2012 main survey.

Question 1, shown in Figure 1.4, calls for a straight-forward reading of data from the graph to answer a question about the context. Students needed to orient themselves to the information presented, identify which data series represents sales for the specified band, which bar represents the specified month within that series, and read the value 500 CDs directly from the vertical axis. The text is simple and clear, creating a very low *communication* demand. The *strategy* required is straightforward: simply to find the specified information in the graph. The *mathematising* demand is to make an inference about the sales situation directly from the graphical model. The *representation* capability is called on at a low level, involving reading a value directly from the graph. The graph format would be familiar to most 15-year-olds, and with effort required only to read the labels to identify what is represented. One axis of the graph is a category axis (months) and the height of the relevant bar is labelled (500) so no understanding of scale is required. The *technical* knowledge required is minimal beyond familiarity with the graph form; and only a direct inference is required, hence very low level demand for *reasoning and argument*. This was an extremely easy item, with some 87% of students identifying the correct response, B.

Question 2 is only slightly more difficult, with about 78% correctly identifying response C. To answer this question, students must observe the relationship between two data series displayed in the bar chart, taking notice of how that relationship changes over the time period shown, in order to recognise that the condition specified in the question was first met in April.

The *communication* demand is similar to that for Question 1. The *strategy* needed is slightly more involved, since multiple elements of the two data series need to be drawn together. The *mathematisation* required again involves making an inference about the sales situation fairly directly from the graph. The *representation* demand is slightly raised from the requirement to read a single data point in Question 1, involving the linking of two data series and the time variable. The demand for *using symbolic, formal and technical language and operations* remains low as only a qualitative comparison is required; and the *reasoning and argument* demand is slightly elevated since a small sequence of reasoning steps is required.

Question 3 is somewhat different from the first two, in that the main focus is on understanding a mathematical relationship depicted in the graph, and extrapolating that relationship to predict the next monthly value. The link to the context is still there, but the main demand is to work with the mathematical information shown. One way to do this would be to read the monthly data values for the series in question, estimate a reasonable average value by which each monthly value is reducing, and apply that same reduction to the data value given for the final month shown. The *communication* demand remains low. The main challenge is to avoid the distraction of the data series of other bands. However, the only common wrong answer was perhaps due to an error in understanding the phrase "the same negative trend". Overall, 15% of students answered C, estimating the sales for July to be equal to the sales for June. They may have chosen the constant value because it maintained the same bad June sales figures into July. The *strategy* needed is clearly more involved than in the first two questions and its implementation requires some monitoring. There are decisions to make, such as whether to use all five February to June data points for this band, or to use the average change from February to June, and whether to calculate exactly, to draw or visualise a trend line or to work with broad estimates noting that each month the sales drop by just over one vertical scale division. The *mathematisation* demand involves a small manipulation of the given model in relation to the context; some calculation is required (repeated subtraction of multi-digit numbers, scale reading between labelled points) that would add to the demand for *using symbolic, formal and technical language and operations*. The *representation* demand involves inferring a trend relationship depicted in the graph; and a small sequence of *reasoning* steps is required to solve the problem. Nevertheless, this item is also relatively easy, with some 76% of students selecting the correct response B in the PISA 2012 main survey administration.

■ Figure 1.5 ■
Items for the unit *CLIMBING MOUNT FUJI*

CLIMBING MOUNT FUJI

Mount Fuji is a famous dormant volcano in Japan.

QUESTION 1

Mount Fuji is only open to the public for climbing from 1 July to 27 August each year. About 200 000 people climb Mount Fuji during this time.

On average, about how many people climb Mount Fuji each day?

A. 340

B. 710

C. 3 400

D. 7 100

E. 7 400

QUESTION 2

The Gotemba walking trail up Mount Fuji is about 9 kilometres (km) long.

Walkers need to return from the 18 km walk by 8 pm.

Toshi estimates that he can walk up the mountain at 1.5 kilometres per hour on average, and down at twice that speed. These speeds take into account meal breaks and rest times.

Using Toshi's estimated speeds, what is the latest time he can begin his walk so that he can return by 8 pm?

....

QUESTION 3

Toshi wore a pedometer to count his steps on his walk along the Gotemba trail.

His pedometer showed that he walked 22 500 steps on the way up.

Estimate Toshi's average step length for his walk up the 9 km Gotemba trail. Give your answer in centimetres (cm).

Answer .... cm

CLIMBING MOUNT FUJI

A second illustrative unit is titled *CLIMBING MOUNT FUJI*, shown in Figure 1.5. The first question is a simple multiple choice item and the second and third questions are constructed response items requiring numerical answers. The third item has partial credit available. This is used for a small proportion of PISA items where qualitatively different kinds of response can be given, and where markedly different abilities can be associated with different kinds of responses.

CLIMBING MOUNT FUJI was used in the PISA 2012 main survey, and then released into the public domain. Questions 1 and 3 lie in the *quantity* content category, since they ask students to calculate with dates and measurements and make conversions. Question 2 has speed as its central concept and is therefore in the *change and relationships* content category.

They each lie in the *societal* context category, since the data relate to information about public access to Mount Fuji and its trails. The first two questions are examples of the *formulating situations mathematically* process category, since the main demand of these questions involves creating a mathematical model that can answer the posed questions.

Question 3 is placed in the *employing mathematical facts, concepts, procedures and reasoning* category because the main demand here is to calculate an average, taking care to convert units appropriately, hence working essentially within the mathematical details of the problem rather than connecting those details with the contextual elements. The three questions were of varying difficulty in the PISA 2012 main survey. Question 1 was of medium difficulty, and Questions 2 and 3 were both very difficult.

Question 1 calls for calculation of the average number of people per day. The text is simple and clear, creating a low *communication* demand. The *strategy* required is of moderate demand, because it involves finding the number of days from the dates provided and using this to find the average. This multiple step solution requires some monitoring, which is also part of the *devising strategies* demand. The *mathematising* demand is very low, because the mathematical quantities required are directly given in the question (number of people per day). Demand for the *representation* capability is similarly low – only numerical information and text are involved. The *technical* knowledge required includes knowing how to find an average, being able to calculate number of days from dates, being able to perform the division (using calculator or not, depending on country assessment policy), and rounding the result appropriately. There is low level demand for *reasoning and argument*. This was an item of medium difficulty, with some 46% of students in the PISA 2012 main survey administration identifying the correct response, C. The two most popular wrong choices were E (which is obtained by using 27 days instead of 31+27 days) with 19% of responses; and A (a place value error) with 12% of responses.

Question 2 is considerably more difficult, with about 12% correct in the PISA 2012 main survey. One factor in this difficulty is that it is a *constructed response* item, rather than *selected response*, so students are given no guidance regarding possible answers, but there are many other factors. About 61% of responses to the PISA 2012 survey administration of this question were wrong answers, not missing.

The *communication* demand is low and, in its receptive aspects, similar to that for Question 1. The constructive communication only requires a numerical response. The *strategy* needed is much more involved, since a plan with three main parts needs to be assembled. Times up and down the mountain need to be calculated from the average speeds, and then the starting time needs to be calculated from the finishing time and the time the walk takes. The *mathematisation* required is moderately high, involving aspects such as understanding how meal times are already included and even that the trail will first be up and then separately down. The *representation* demand is minimal, with only the interpretation of text required. The demand for *using symbolic, formal and technical language and operations* is moderately high: all the calculations are relatively simple (although division by the decimal 1.5 km per hour may be challenging) but it requires sustained accuracy, and the formula for time from speed and distance is required either implicitly or explicitly. The *reasoning and argument* demand is also moderately high.

Question 3 is also quite difficult. The main focus is to calculate average step length from distance and number of steps, with unit conversions required. For this item, 11% of responses in the PISA 2012 survey administration gained full credit for the correct response 40 cm, and a further 4% gained partial credit for responses such as 0.4 (the answer left in metres) or 4 000 where an incorrect conversion factor from metres to centimetres has probably been used. For the PISA 2012 main survey administration, 62% of responses were incorrect answers, not missing. The *communication* demand remains low as with the earlier questions, since the text is fairly clear and easy to interpret and the requirement for a single number as a sufficient answer. The *strategy* needed for Question 3 is similar to that for Question 1 – both require finding an average. Although both use similar models to find 'averages', the *reasoning and argument* needed for Question 3 is more involved than for Question 1. In Question 1, the quantity required is "people per day" where the number of people is given and the number of days is readily calculated. Question 3 requires "step length" to be calculated from a total distance and a total number of steps. More reasoning is required to link these quantities in Question 3 (for example linking the given distance with the length). The *mathematisation* demand is similarly higher in Question 3, understanding how the real world quantity of step length relates to the overall measures. An appreciation of the real world context, including that step length is likely to be around 50 cm (rather than 500 cm or 0.5 cm), is also useful for monitoring the reasonableness of the answer. The demand for *using symbolic, formal and technical language and operations* is moderately high, because of the division of a small number (9 km) by a large number (22 500 steps) and the need for using known conversion factors. The *representation* demand is again low, since only text is involved.

PIZZAS

The open constructed-response item *PIZZAS* shown in Figure 1.6 is simple in form, yet rich in content, and illustrates various elements of the mathematics framework. It was initially used in the first PISA field trial in 1999, then was released for illustrative purposes and has appeared as a sample item in each version of the PISA mathematics framework published since 2003. This was one of the most difficult items used in the 1999 field trial item pool, with only 11% correct.

PIZZAS is set in a *personal* context with which many 15-year-olds would be familiar. The context category is personal since the question posed is which pizza provides the purchaser with the better value for the money. It presents a relatively low reading demand, thereby ensuring the efforts of the reader can be directed almost entirely to the underlying mathematical intentions of the task.

■ Figure 1.6 ■

Item for the unit *PIZZAS*

PIZZAS

A pizzeria serves two round pizzas of the same thickness in different sizes. The smaller one has a diameter of 30 cm and costs 30 zeds. The larger one has a diameter of 40 cm and costs 40 zeds.

Which pizza is better value for money? Show your reasoning.

....

Every day terms from the real world must be interpreted mathematically (*round, same thickness, different sizes*). The size variable is given mathematical definition in the diameters provided for the two pizzas. The costs are provided in the neutral currency *zeds*. Size and cost are linked through the concept of *value for money*.

The item draws on several areas of mathematics. It has geometrical elements that would normally be classified as part of the *space and shape* content category. The pizzas can be modelled as thin circular cylinders, so the area of a circle is needed. The question also involves the *quantity* content category with the implicit need to compare the quantity of pizza to amount of money. However, the key to this problem lies in the conceptualisation of the relationships among properties of the pizzas, and how the relevant properties change from the smaller pizza to the larger one. Because those aspects are at the heart of the problem, this item is categorised as belonging to the *change and relationships* content category.

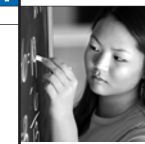
The item belongs to the *formulating* process category. A key step to solving this problem, indeed the major cognitive demand, is to formulate a mathematical model that encapsulates the concept of *value for money*. The problem solver must recognise that because pizzas ideally have uniform thickness and the thicknesses are the same, the focus of analysis can be on the area of the circular surface of the pizza instead of volume or mass. The relationship between amount of pizza and amount of money is then captured in the concept of *value for money* modelled as 'cost per unit of area'. Variations such as area per unit cost are also possible. Within the mathematical world, *value for money* can then be calculated directly and compared for the two circles, and is a smaller quantity for the larger circle. The real world interpretation is that the larger pizza represents better value for money.

An alternative form of reasoning, which reveals even more clearly the item's classification in *change and relationships*, would be to say (explicitly or implicitly) that the area of a circle increases in proportion to the square of the diameter, so has increased in the ratio of $(4/3)^2$, while the cost has only increased in the proportion of (4/3). Since $(4/3)^2$ is greater than (4/3), the larger pizza is better value.

While the primary demand and the key to solving this problem comes from formulating, placing this item in the *formulating situations mathematically* process category, aspects of the other two mathematical process are also apparent in this item. The mathematical model, once formulated, must then be employed effectively, with the application of appropriate reasoning along with the use of appropriate mathematical knowledge and area and rate calculations. The result must then be interpreted properly in relation to the original question.

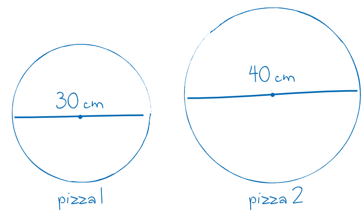
The solution process for *PIZZAS* demands the activation of the fundamental mathematical capabilities to varying degrees. *Communication* comes in to play at a relatively low level in reading and interpreting the rather straight-forward text of the problem, and is called on at a higher level with the need to present and explain the solution. The need to *mathematise* the situation is a key demand of the problem, specifically the need to formulate a model that captures *value for money*. The problem solver must devise a *representation* of relevant aspects of the problem, including the symbolic representation of the formula for calculating area, and the expression of rates that represent *value for money*, in order to develop a solution. The *reasoning* demands (for example, to decide that the thickness can be ignored, and justifying the approach taken and the results obtained) are significant, and the need for *devising strategies* to control the calculation and modelling processes required is also a notable demand for this problem. *Using symbolic, formal and technical language and operations* comes into play with the conceptual, factual and procedural knowledge required to process the circle geometry, and the calculations of the rates. *Using mathematical tools* is evident at a relatively low level if students use a calculator efficiently.

In Figure 1.7, a sample student response to the *PIZZAS* item is presented, to further illustrate the framework constructs. A response like this would be awarded full credit.



■ Figure 1.7 ■

Sample response to *PIZZAS*



An important part of formulation	The thickness is the same so I can compare areas.
Employing knowledge from Space and shape, and Quantity	Area of pizza 1 $= \pi r^2$ $= \pi \times 15 \times 15 \text{ cm}^2$ $= 706{,}5 \text{ cm}^2$ Area of pizza 2 $= \pi r^2$ $= \pi \times 20 \times 20 \text{ cm}^2$ $= 1256 \text{ cm}^2$
Formulating a mathematical model to measure value for money	Cost per cm² of pizza 1 = 30 zeds/706,5 cm² = 0.04 zeds/cm² Cost per cm² of pizza 2 = 40 zeds/1256 cm² = 0.03 zeds/cm²
Interpreting mathematical result in real world terms	So pizza 2 is cheaper per cm² + is better value

LITTER

The item *LITTER* shown in Figure 1.8 is also presented to illustrate aspects of the mathematics framework. This constructed-response item was used in the PISA 2003 main survey and then released into the public domain. The average percent correct for this item in OECD countries was slightly over 51%, placing it near to the middle of the item pool in difficulty.

■ Figure 1.8 ■

Item for the unit *LITTER*

For a homework assignment on the environment, students collected information on the decomposition time of several types of litter that people throw away:

Type of litter	Decomposition time
Banana peel	1-3 years
Orange peel	1-3 years
Cardboard boxes	0.5 years
Chewing gum	20-25 years
Newspapers	A few days
Polystyrene cups	Over 100 years

A student thinks of displaying the results in a bar graph.

*Give **one** reason why a bar graph is unsuitable for displaying these data.*

....

This item is set in a *scientific* context, since it deals with data of a scientific nature (decomposition time). The mathematical content category is *uncertainty and data*, since it primarily relates to the interpretation and presentation of data, although *quantity* is involved in the implicit demand to appreciate the relative sizes of the time intervals involved. The mathematical process category is *interpreting, applying and evaluating mathematical outcomes* since the focus is on evaluating the effectiveness of the mathematical outcome (in this case an imagined or sketched bar graph) in portraying the data about the real world contextual elements. The item involves reasoning about the data presented, thinking

mathematically about the relationship between the data and their presentation, and evaluating the result. The problem solver must recognise that these data would be difficult to present well in a bar graph for one of two reasons: either because of the wide range of decomposition times for some categories of litter (this range cannot readily be displayed on a standard bar graph), or because of the extreme variation in the time variable across the litter types (so that on a time axis that allows for the longest period, the shortest periods would be invisible). Student responses such as those reproduced below have been awarded credit for this item.

RESPONSE 1

"Because it would be hard to do in a bar graph because there are 1-3, 1-3, 0.5, etc. so it would be hard to do it exactly."

RESPONSE 2

"Because there is a large difference from the highest sum to the lowest therefore it would be hard to be accurate with 100 years and a few days."

The solution process for *LITTER* demands the activation of the fundamental mathematical capabilities as follows. *Communication* comes in to play with the need to read the text and interpret the table, and is also called on at a higher level with the need to answer with brief written reasoning. The demand to *mathematise* the situation arises at a low level with the need to identify and extract key mathematical characteristics of a bar graph as each type of litter is considered. The problem solver must interpret a simple tabular *representation* of data, and must imagine a graphical representation, and linking these two representations is a key demand of the item. The *reasoning* demands of the problem are at a relatively low level, as is the need for *devising strategies*. *Using symbolic, formal and technical language and operations* comes into play with the procedural and factual knowledge required to imagine construction of bar graphs or to make a quick sketch, and particularly with the understanding of scale needed to imagine the vertical axis. *Using mathematical tools* is likely not needed.

ROCK CONCERT

A further illustrative item, *ROCK CONCERT*, is presented in Figure 1.9. This selected-response item (here simple multiple choice) was used in the field trial prior to the PISA 2003 survey, then was released into the public domain for illustrative purposes. About 28% of sampled students got this item correct (choice C), making it a moderately difficult item relative to the pool of items used in the field trial. *ROCK CONCERT* is set in a societal context, because the item is set at the level of the rock concert organisation, even though it draws on personal experience of being in crowds. It is classified within the *quantity* content category because of the numerical calculation required, though it also has some elements that relate to the *space and shape* category.

■ Figure 1.9 ■

Item for the unit *ROCK CONCERT*

For a rock concert, a rectangular field of size 100 m by 50 m was reserved for the audience. The concert was completely sold out and the field was full with all the fans standing.

Which one of the following is likely to be the best estimate of the total number of people attending the concert?

A. 2 000

B. 5 000

C. 20 000

D. 50 000

E. 100 000

This item calls on each of the three process categories but the primary demand comes from *formulating situations mathematically*, with the need to make sense of the contextual information provided (the field size and shape; the rock concert is full; fans are standing) and translate it into a useful mathematical form. There is also the need to identify information that is missing, but that could reasonably be estimated based on real-life knowledge and assumptions. Specifically there is a need to devise a model for the space required for an individual fan or a group of fans. Working within mathematics, the problem solver needs to *employ mathematical concepts, facts, procedures and reasoning* to link the area of the field and the area occupied by a fan to the number of fans, making the quantitative comparisons needed. And *interpreting, applying and evaluating mathematical outcomes* is required to check the reasonableness of the solution, or to evaluate the answer options against the mathematical results of calculations performed.

An alternative model is to imagine the fans standing uniformly in equal rows across the field and to estimate the number of fans by multiplying the estimated number of rows by the estimated number of fans in each row. Problem solvers with strong skills in formulating mathematical models may appreciate the effectiveness of this rows-and-columns model, despite the stark contrast between it and the behaviour of fans at a rock concert. The correct answer is insensitive to which of several reasonable models is being adopted by the student.

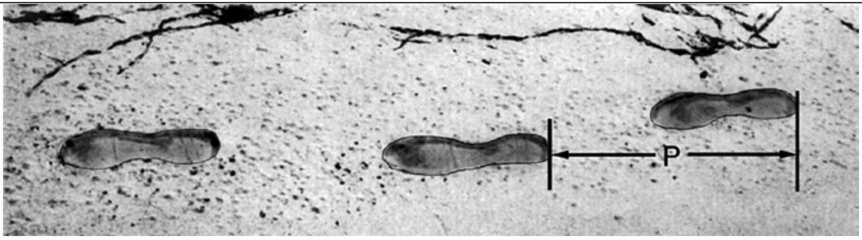
The fundamental mathematical capabilities come into play for this question in the following ways. *Communication* is called on at a relatively low level with the need to read and understand the text. The mathematical importance of words such as *rectangular and size*, the phrase the *field was full*, and the instruction to *estimate*, must all be interpreted and understood. Some real-world knowledge will help to do this. The task has a significant *mathematisation* demand, since solving the problem would require making certain assumptions about the space that a person might occupy while standing as well as requiring the creation of a basic model such as (number of fans) x (average space for a fan) = (area of field). To do this one must *represent* the situation mentally or diagrammatically, as part of formulating the model to link the space for a fan with the area of the field. *Devising a strategy* comes into the process of solving this problem at several stages, such as when deciding on how the problem should be approached, when imagining what kind of model could be useful to capture the space occupied by a fan at the concert, and when recognising the need for some checking and validation procedures. One solution strategy would involve postulating an area for each person, multiplying it by the number of people given in each of the options provided, and comparing the result to the conditions given in the question. Alternatively, the reverse could be done, starting with the area provided and working backwards using each of the response options to calculate the corresponding space per person, and deciding which one best fits the criteria established in the question. *Using symbolic, formal and technical language and operations* comes into play in implementing whatever strategy was adopted, by interpreting and using the dimensions provided and in carrying out the calculations required to relate the field area to the area for an individual. *Reasoning and argument* would come in to play with the need to think clearly about the relationship between the model devised, the resulting solution, and the real context, in order to validate the model used and to check that the correct answer is chosen. *Using mathematical tools* is unlikely to be needed.

WALKING

The PISA unit *WALKING*, presented in Figure 1.10, shows a somewhat counter-intuitive but well-established algebraic relationship between two variables, based on the observation of a large number of men walking at a natural pace, and asks students two questions that demand activation of algebraic knowledge and skills. For the second question, strategic thinking, reasoning and argument capabilities are also demanded at a level that challenges many 15-year-olds. These items were used in the PISA 2003 main survey, then released into the public domain and have subsequently been used as illustrative items in the PISA 2009 framework and in other publications. Both questions require students to work with the information given and to construct their response. Both items fit within the same framework categories: the *change and relationships* content category, since they relate to the relationships among variables, in this case expressed in algebraic form; the *personal* context category, since they focus on matters relating directly to the experience and perspective of the individual; and the *employing mathematical facts, concepts, procedures and reasoning* process category, since the problems have been expressed in terms that already have mathematical structure, and the work required is largely intra-mathematical manipulation of mathematical concepts and objects.

■ Figure 1.10 ■

Items for the unit *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 $\frac{n}{P} = 140$ gives an approximate relationship between n and P where:

n = number of steps per minute, and

P = pacelength in metres.

QUESTION 1

If the formula applies to Heiko's walking and Heiko takes 70 steps per minute, what is Heiko's pacelength? Show your work.

QUESTION 2

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking.

Calculate Bernard's walking speed in metres per minute and in kilometres per hour. Show your working out.

Question 1 had an international percentage correct figure of 36% in the 2003 main survey, making it more difficult than about 70% of items in the 2003 pool. This is surprising, since mathematically all that is required is to substitute the value $n=70$ into the formula, and implement some reasonably straight-forward algebraic manipulation of the formula to find the value of P. This item illustrates the observation that has frequently been made about PISA survey items that when test questions are placed in some real world context, even when the mathematical components are presented clearly in the question, 15-year-old students often struggle to apply their mathematical knowledge and skills effectively.

The fundamental mathematical capabilities come into play for this question in the following ways. *Communication* is called on with the need to read and understand the stimulus, and later to articulate a solution and show the work involved. The task has no real *mathematisation* demand, since a mathematical model is provided in a form that would be familiar to many 15-year-old students. The *representation* demand is significant, given that the stimulus includes a graphic element, text and an algebraic expression that must be related to each other. *Devising a strategy* comes into the solution process at a very low level, since the strategy needed is very clearly expressed in the question. Minimal *reasoning and argument* is needed, again because the task is clearly stated and all required elements are obvious. *Using symbolic, formal and technical language and operations* comes into play in performing the substitution and manipulating the expression to make P the subject of the equation.

Question 2 is more difficult, with an international average percentage correct of 20%, meaning it was among the most difficult 10% of items used in the 2003 PISA survey. *Devising a strategy* for this question is complex because of the number of steps involved, and the resulting need to keep focused on the desired endpoint: P is known and so n can be found from the given equation; multiplying n by P gives the speed in number of metres travelled per minute; then proportional reasoning can be used to change the units of speed to kilometres per hour. Three levels of credit were available to accommodate solutions for which only partial progress towards a complete solution was achieved. The difference in the observed percentage correct for Question 2 compared to Question 1 can probably best be explained by describing the different activation of the fundamental mathematical capabilities that are required. The *communication* required for the two questions is comparable at the stage of reading and understanding the question, but in Question 2 the diagram has to be used to explicitly link one step and the given pacelength, a relationship not needed in Question 1, and the presentation of the solution demands higher level expressive communication skills for Question 2. The task has a new *mathematisation* demand, since solving the problem would require devising a proportional model for Bernard's walking speed in the units requested. Such a solution process requires activation of effective and sustained control mechanisms across a multi-step procedure, hence the *devising a strategy* capability is required at a much higher level than was the case for Question 1. The *representation* demands in the second question go beyond those needed for Question 1 with the need to work more actively with the given algebraic representation. Implementing the strategy devised and using the representations identified involves *using symbolic, formal and technical language and operations* that includes the algebraic manipulations, and the application of proportions and arithmetic calculations to perform the required conversions. *Reasoning and argument* comes in to play throughout with the sustained and connected thought processes required to proceed with the solution. *Using mathematical tools* is evident at a relatively low level if students use a calculator efficiently.

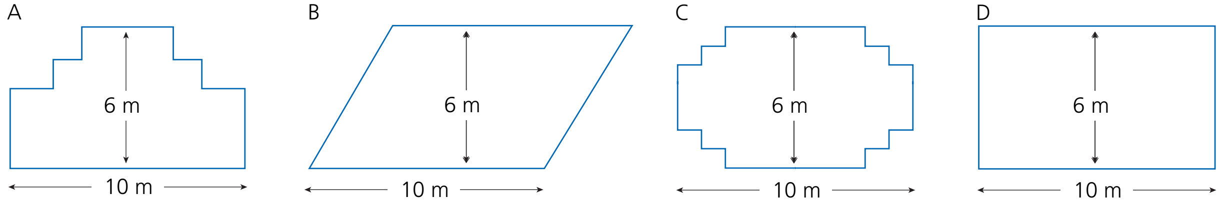
CARPENTER

The PISA item *CARPENTER* is presented in Figure 1.11. This item was used in both the PISA 2000 and 2003 surveys, and then released into the public domain. It illustrates a form of selected-response item known as the complex multiple-choice format, for which students must select one response from options attached to each of a number of statements or questions. In this case, students gained full credit by correctly identifying that all designs except Design B can be made with the specified amount of timber.

The item fits into the *space and shape* content category, since it deals with properties of shapes. It is associated with the *occupational* context category as it deals with a work task of a carpenter. The item is classified under the *employing mathematical concepts, facts, procedures and reasoning* process category, since most of the work involves applying procedural knowledge to well defined mathematical objects; although it also involves some degree of *interpreting, applying and evaluating mathematical outcomes* given the need to link the mathematical objects represented to the contextual element – the constraint imposed by the available timber.

■ Figure 1.11 ■
Item for the unit *CARPENTER*

A carpenter has 32 metres of timber and wants to make a border around a garden bed. He is considering the following designs for the garden bed.



Circle either "Yes" or "No" for each design to indicate whether the garden bed can be made with 32 metres of timber.

Garden bed design	Using this design, can the garden bed be made with 32 metres of timber?
Design A	Yes / No
Design B	Yes / No
Design C	Yes / No
Design D	Yes / No

This was one of the more difficult items in the PISA 2003 survey, with a correct response rate of a little less than 20%. It can be solved by the application of geometrical knowledge and reasoning. Enough information is given to enable direct calculation of the exact perimeter for Designs A, C and D, each of which is 32 metres. However, insufficient information is given for Design B; therefore a different approach is required. It can be reasoned that while the 'horizontal' components of the four shapes are equivalent, the oblique sides of Design B are longer than the sum of the 'vertical' components of each of the other shapes.

The *communication* capability is called on in reading and understanding the question, and to link the information provided in the text with the graphical *representation* of the four garden beds. The task has been presented in overtly mathematical form, hence no *mathematisation* is needed. Real world considerations, such as the lengths of the pieces of timber available and the geometry of the corners, do not come into the problems as posed here. The key capability demanded to solve the problem is the *reasoning and argument* needed to identify Design B which has too great a perimeter, and to appreciate that the lengths of the 'vertical' components of Design A are in themselves unknown, but that the total 'vertical' length is known (similarly with Design C with both vertical and horizontal lengths). *Devising a strategy* involves recognising that the perimeter information needed can be found in spite of the fact that some of the individual lengths are not known. *Using symbolic, formal and technical language and operations* is needed in the form of an understanding and manipulating of the perimeter of the shapes presented, including both the properties of the sides, and the addition of the side lengths. *Using mathematical* tools is likely not needed.

Notes

1. In some countries, "mathematical tools" can also refer to established mathematical procedures such as algorithms. For the purposes of the PISA framework, "mathematical tools" refers only to the physical and digital tools described in this section.

2. The standards for two sets of countries were analysed. The sets were nine OECD countries (Australia [New South Wales], [Flemish] Belgium, Canada [Alberta], Finland, Ireland, Japan, Korea, New Zealand and the United Kingdom), and six high-performing countries ([Flemish] Belgium, Canada [Alberta], Chinese Taipei, Finland, Korea, and Singapore). A constraint of the analysis was that standards had to be available in English.

3. Those familiar with earlier frameworks will note that the category *uncertainty* is now called *uncertainty and data*. This name change is intended only to describe the category more clearly, and is not a fundamental change to the category itself.

4. In 2006, PISA pilot tested a computer-based science assessment, and in 2009 included an optional digital reading assessment.

References

Bennett, R. (2003), *Online Assessment and the Comparability of Score Meaning*, Educational Testing Service, Princeton, New Jersey, *www.ets.org/Media/Research/pdf/RM-03-05-Bennett.pdf*.

Bennett, R.E., J. Braswell, A. Oranje, B. Sandene, B. Kaplan and **F. Yan** (2008), "Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP", *Journal of Technology, Learning, and Assessment*, No. 6, Vol. 9.

Common Core State Standards Initiative (2010), *Common Core State Standards for Mathematics*, Common Core State Standards Initiative, Washington, D.C., *http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf*.

Devlin, K. (1994), *Mathematics: The Science of Patterns: The Search for Order in Life, Mind and the Universe*, W. H. Freeman Scientific American Library, New York.

Hoyles, C., A. Wolf, S. Molyneux-Hodgson and **P. Kent** (2002), *Mathematical skills in the workplace: final report to the Science Technology and Mathematics Council, Project Report*, Institute of Education, University of London, Science, Technology and Mathematics Council, London, *http://eprints.ioe.ac.uk/1565/1/Hoyles2002MathematicalSkills.pdf*.

Mason, B., M. Patry and **D. Berstein** (2001), "An examination of the equivalence between non-adaptive computer based and traditional testing", *Journal of Education Computing Research*, No. 24, Vol. 1, pp. 29-39.

Moore, D. (1997), "New pedagogy and new content: The case of statistics", *International Statistical Review*, No. 65, Vol. 2, pp. 123-137.

National Council of Teachers of Mathematics (2000), *Principles and Standards for School Mathematics*, NCTM, Reston, Virginia, *www.nctm.org/standards/*.

Niss, M. and **T. H. Jensen** (2002), *Kompetencer og matematiklæring: Ideer og inspiration til udvikling af matematikundervisning i Danmark*, Uddannelsesstyrelsens temahæfteserie, No. 18, Ministry of Education, Copenhagen, *http://pub.uvm.dk/2002/kom/*.

Niss, M. (2003), "Mathematical Competencies and the learning of mathematics: The Danish KOM Project", in Gagatsis A. and S. Papastavridis (eds.), *3rd Mediterranean Conference on Mathematics Education*, The Hellenic Mathematical Society and Cyprus Mathematical Society, Athens, pp. 115-124, *http://w3.msi.vxu.se/users/hso/aaa_niss.pdf*.

Niss, M., W. Blum and **P. Galbraith** (2007), "Introduction", in Blum, W., P. Galbraith, H.-W. Henn and M. Niss (eds.), *Modelling and Applications in Mathematics Education (The 14th ICMI Study)*, Springer, New York, pp. 3-32.

Niss, M. and **T. Højgaard** (eds.) (2011), "Competencies and Mathematical Learning: Ideas and inspiration for the development of mathematics teaching and learning in Denmark", Ministry of Education, Report No. 485, Roskilde University, Roskilde, *https://pure.au.dk/portal/files/41669781/THJ11_MN_KOM_in_english.pdf*.

OECD (2003), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, PISA, OECD Publishing.

OECD (2010), *Pathways to Success: How knowledge and skills at age 15 shape future lives in Canada*, PISA, OECD Publishing, *www.oecd.org/dataoecd/59/35/44574748.pdf*.

Qualifications and Curriculum Authority (2007), "Mathematics: Programme of study for key stage 3 and attainment targets", Qualifications and Curriculum Authority, London, *http://media.education.gov.uk/assets/files/pdf/q/mathematics%202007%20programme%20of%20study%20for%20key%20stage%203.pdf*.

Richardson M., J.-A. Baird, J. Ridgway, M. Ripley, D. Shorrocks-Taylor and **M. Swan** (2002), "Challenging Minds? Students' perceptions of computer-based World Class Tests of problem solving", *Computers in Human Behavior*, Vol. 18, Issue 6, November, pp. 633-649.

Sandene, B., N. Horkay, R. Bennett, N. Allen, J. Braswell, B. Kaplan, and **A. Oranje** (2005), *Online Assessment in Mathematics and Writing: Reports From the NAEP Technology Based Assessment Project*, Research and Development Series (NCES 2005–457), U.S. Department of Education, National Center for Education Statistics, Washington, D.C., U.S. Government Printing Office, *www.nces.ed.gov/nationsreportcard/pdf/studies/2005457_1.pdf*.

Stacey, K. and **D. Wiliam** (forthcoming), "Technology and Assessment in Mathematics", in Clements M. A. (Ken), A. Bishop, C. Keitel, J. Kilpatrick and F. Leung (eds.), *Third International Handbook of Mathematics Education*, Springer, pp. 721 - 752.

Steen, L. (1990), *On the Shoulders of Giants: New Approaches to Numeracy*, National Academy Press Washington, D.C.

Thomson, S. and **K. Hillman** (2010), *Against the Odds: Influences on the Post-School Success of 'Low Performers'*, NCVER, Adelaide, *www.ncver.edu.au/publications/2285.html*.

Turner, R. (2012), "Some Drivers of Test Item Difficulty in Mathematics", Paper presented at the Annual Meeting of the American Educational Research Association (AERA), 13-17 April 2012, Vancouver, *http://research.acer.edu.au/pisa/4/*.

Turner, R. and **R.J. Adams** (2012), "Some drivers of test item difficulty in mathematics: an analysis of the competency rubric", Paper presented at the Annual Meeting of the American Educational Research Association (AERA), 13-17 April 2012, Vancouver, *http://research.acer.edu.au/pisa/7/* .

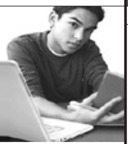
Turner, R., J. Dossey, W. Blum and **M. Niss** (forthcoming), "Using mathematical competencies to predict item difficulty in PISA", in Prenzel, M., M. Kobarg, K. Schöps and S. Rönnebeck (eds.), *Research on PISA: Research Outcomes of the PISA Research Conference 2009*, Springer, New York. pp. 23–27.

Watson, J.M. and **R. Callingham** (2003), "Statistical literacy: A complex hierarchical construct", *Statistics Education Research Journal*, No. 2, Vol. 2, pp. 3-46.

2

PISA 2012 Reading Framework

This chapter discusses the conceptual framework underlying the PISA 2012 assessment of students' reading competencies. It provides the PISA definition of reading literacy and presents the elements of the survey which have remained consistent throughout the previous cycles, along with a new element introduced in PISA 2009: reading and understanding digital texts. It describes how PISA assesses and analyses print and digital reading tasks, as well as the way in which students navigate through digital texts and respond to the format of tasks. Sample print and digital reading items are included throughout the chapter to further illustrate how students' skills are measured.



INTRODUCTION

Reading literacy was the major domain assessed in 2000, for the first PISA cycle (PISA 2000) and in 2009, for the fourth PISA cycle (PISA 2009). For the fifth PISA cycle (PISA 2012), reading is a minor domain and its framework has not changed from the previous cycle, PISA 2009 (OECD, 2009). There were two major modifications to the PISA 2009 version of the reading framework: the incorporation of the reading of digital texts and the elaboration of the constructs of reading engagement and metacognition.

Proficiency in reading literacy is a key to unlocking not only the world of printed texts but also digital texts, which are becoming an increasingly important part of students' and adults' reading. In all countries, Internet use is closely linked with socio-economic status and education (Sweets and Meates, 2004). Yet the requirement to use computers is not confined to particular social and economic strata. Beyond the workplace, computer technology has a growing importance in personal, social and civic life (Pew Internet and American Life Project, 2005).

While many of the skills required for print and digital reading are similar, digital reading demands that new emphases and strategies be added to the repertoires of readers. Gathering information on the Internet requires skimming and scanning through large amounts of material and immediately evaluating its credibility. Critical thinking, therefore, has become more important than ever in reading literacy (Halpern, 1989; Shetzer and Warschauer, 2000; Warschauer, 1999). Warschauer concludes that overcoming the "digital divide" is not only a matter of achieving online access, but also of enhancing people's abilities to integrate, evaluate and communicate information.

The new demands on reading proficiency created by the digital world led to the framework's inclusion of digital reading in the PISA 2009 assessment, acknowledging the fact that any definition of reading in the 21st century needs to encompass both printed and digital texts. An assessment of digital reading was also included in PISA 2012. Not all participating countries elected to take part in the administration of the digital reading assessment either in PISA 2009 or in PISA 2012, which was therefore implemented as an international option. Twenty-three OECD countries and nine partner countries and economies chose this option in PISA 2012, an increase of more than 50% over the PISA 2009 numbers.

Changes in our concept of reading since 2000 have already led to an expanded definition of reading literacy, which recognises motivational and behavioural characteristics of reading alongside cognitive characteristics. In light of recent research, reading engagement and metacognition were featured more prominently in the PISA 2009 reading framework as elements that can make an important contribution to policy makers' understanding of factors that can be developed, shaped and fostered as components of reading literacy. However, in PISA 2012, reading is a minor domain and no data on engagement or metacognition in reading were collected.

The PISA framework for assessing the reading literacy of students towards the end of compulsory education, therefore, must focus on reading literacy skills that include finding, selecting, interpreting and evaluating information from the full range of texts associated with situations in the classroom and also those that reach beyond the classroom.

This chapter discusses the conceptual framework underlying the PISA 2012 assessment of students' reading competencies. The definition of the domain is the same as in PISA 2009 when it was, for the second time, the major domain assessed, apart from a new element: reading and understanding digital texts. It describes how PISA assesses and analyses digital reading tasks, and the way in which students navigate through texts and respond to the format of tasks. Sample print and digital reading items are included throughout the chapter to further illustrate how students' skills are measured.

DEFINING READING LITERACY

Definitions of reading and reading literacy have changed over time in parallel with changes in society, economy, and culture. The concept of learning, particularly the concept of lifelong learning, has expanded the perception of reading literacy. Literacy is no longer considered to be an ability acquired only in childhood during the early years of schooling. Instead, it is viewed as an expanding set of knowledge, skills and strategies that individuals build on throughout life in various contexts, through interaction with their peers and the wider community.

Cognitive-based theories of reading literacy emphasise the interactive nature of reading and the constructive nature of comprehension, in the print medium (Binkley and Linnakylä, 1997; Bruner, 1990; Dole et al., 1991) and to an even greater extent in the digital medium (Fastrez, 2001; Legros and Crinon, 2002; Leu, 2007; Reinking, 1994). The reader generates meaning in response to text by using previous knowledge and a range of text and situational cues that are often socially and culturally derived. While constructing meaning, the reader uses various processes, skills and strategies to foster, monitor and maintain understanding. These processes and strategies are expected to vary with context and purpose as readers interact with a variety of continuous and non-continuous texts in the print medium and (typically) with multiple texts in the digital medium.

The PISA 2012 definition of reading literacy is as follows:

> *Reading literacy is 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.*

Reading literacy…

The term "reading literacy" is preferred to "reading" because it is likely to convey to a non-expert audience more precisely what the survey is measuring. "Reading" is often understood as simply decoding, or even reading aloud, whereas the intention of this survey is to measure something broader and deeper. Reading literacy includes a wide range of cognitive competencies, from basic decoding, to knowledge of words, grammar and larger linguistic and textual structures and features, to knowledge about the world.

In this study, "reading literacy" is intended to express the active, purposeful and functional application of reading in a range of situations and for various purposes. According to Holloway (1999), reading skills are essential to the academic achievement of middle- and high school students. PISA assesses a wide range of students. Some will go on to university; some will pursue further studies in preparation for joining the labour force; some will enter the workforce directly after completing compulsory education. Achievement in reading literacy is not only a foundation for achievement in other subject areas within the education system, but also a prerequisite for successful participation in most areas of adult life (Cunningham and Stanovich, 1998; Smith et al., 2000). Indeed, regardless of their academic or labour-force aspirations, students' reading literacy is important for their active participation in their community and economic and personal life.

Reading literacy skills matter not just for individuals, but for economies as a whole. Policy makers and others are coming to recognise that in modern societies, human capital – the sum of what the individuals in an economy know and can do – may be the most important form of capital. Economists have for many years developed models showing generally that a country's education levels are a predictor of its economic growth potential (Coulombe et al., 2004).

…is understanding, using, reflecting on…

The word "understanding" is readily connected with "reading comprehension", a well-accepted element of reading. The word "using" refers to the notions of application and function – doing something with what we read. "Reflecting on" is added to "understanding" and "using" to emphasise the notion that reading is interactive: readers draw on their own thoughts and experiences when engaging with a text. Of course, every act of reading requires some reflection, drawing on information from outside the text. Even at the earliest stages, readers draw on symbolic knowledge to decode a text and require a knowledge of vocabulary to construct meaning. As readers develop their stores of information, experience and beliefs, they constantly, often unconsciously, test what they read against outside knowledge, thereby continually reviewing and revising their sense of the text.

…and engaging with…

A reading literate person not only has the skills and knowledge to read well, but also values and uses reading for a variety of purposes. It is therefore a goal of education to cultivate not only proficiency but also engagement in reading. Engagement in this context implies the motivation to read and comprises a cluster of affective and behavioural characteristics that include an interest in and enjoyment of reading, a sense of control over what one reads, involvement in the social dimension of reading, and diverse and frequent reading practices.

…written texts…

The term "written texts" is meant to include all those coherent texts in which language is used in its graphic form, whether printed and digital. Instead of the word "information", which is used in some other definitions of reading, the term "texts" was chosen because of its association with written language and because it more readily connotes literary as well as information-focused reading.

These texts do not include aural language artefacts such as voice recordings; nor do they include film, TV, animated visuals, or pictures without words. They do include visual displays such as diagrams, pictures, maps, tables, graphs and comic strips that include some written language (for example, captions). These visual texts can exist either independently or they can be embedded in larger texts. Digital texts are distinguished from printed texts in a number of respects, including physical readability; the amount of text visible to the reader at any one time; the way different parts of a text and different texts are connected with one another through hypertext links; and, given these text characteristics, the way

that readers typically engage with digital texts. To a much greater extent than with printed or hand-written texts, readers need to construct their own pathways to complete any reading activity associated with a digital text.

…in order to achieve one's goals, develop one's knowledge and potential, and participate in society.

This phrase is meant to capture the full scope of situations in which reading literacy plays a role, from private to public, from school to work, from formal education to lifelong learning and active citizenship. "To achieve one's goals and to develop one's knowledge and potential" spells out the idea that reading literacy enables the fulfilment of individual aspirations – both defined ones, such as graduating or getting a job, and those less defined and less immediate that enrich and extend personal life and lifelong education. The word "participate" is used because it implies that reading literacy allows people to contribute to society as well as to meet their own needs. "Participating" includes social, cultural and political engagement.

ORGANISING THE DOMAIN

This section describes how the domain is represented, a vital issue because the organisation and representation of the domain determines the test design and, ultimately, the evidence about student proficiencies that can be collected and reported.[1]

Reading is a multidimensional domain. While many elements are part of the construct, not all can be taken into account in building the PISA assessment. Only those considered most important were selected.

The PISA reading literacy assessment is built on three major task characteristics to ensure a broad coverage of the domain:

- *situation*, which refers to the range of broad contexts or purposes for which reading takes place;
- *text*, which refers to the range of material that is read; and
- *aspect*, which refers to the cognitive approach that determines how readers engage with a text.

In PISA, features of the text and aspect variables (but not of the situation variable) are also manipulated to influence the difficulty of a task.

Reading is a complex activity. The elements of reading, do not exist independently of one another in neat compartments. The assignment of texts and tasks to framework categories does not imply that the categories are strictly partitioned or that the materials exist in atomised cells determined by a theoretical structure. The framework scheme is provided to ensure coverage, to guide the development of the assessment and to set parameters for reporting, based on what are considered the marked features of each task.

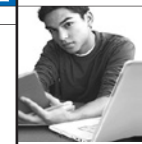
Situation

The PISA situation variables were adapted from the Common European Framework of Reference (CEFR) developed for the Council of Europe (Council of Europe, 1996). The four situation variables – personal, public, educational and occupational – are described in the following paragraphs.

The *personal* situation relates to texts that are intended to satisfy an individual's personal interests, both practical and intellectual. This category also includes texts that are intended to maintain or develop personal connections with other people. It includes personal letters, fiction, biography, and informational texts that are intended to be read to satisfy curiosity, as a part of leisure or recreational activities. In the digital medium it includes personal e-mails, instant messages and diary-style blogs.

The *public* category describes the reading of texts that relate to activities and concerns of the larger society. The category includes official documents and information about public events. In general, the texts associated with this category assume a more or less anonymous contact with others; they also therefore include forum-style blogs, news websites and public notices that are encountered both on line and in print.

The content of *educational* texts is usually designed specifically for the purpose of instruction. Printed text books and interactive learning software are typical examples of material generated for this kind of reading. Educational reading normally involves acquiring information as part of a larger learning task. The materials are often not chosen by the reader, but instead assigned by an instructor. The model tasks are those usually identified as "reading to learn" (Sticht, 1975; Stiggins, 1982).



Many 15-year-olds will move from school into the labour force within one to two years. A typical *occupational* reading task is one that involves the accomplishment of some immediate task. It might include searching for a job, either in a print newspaper's classified advertisement section, or on line; or following workplace directions. The model tasks of this type are often referred to as "reading to do" (Sticht, 1975; Stiggins, 1982).

Situation is used in PISA reading literacy to define texts and their associated tasks, and refers to the contexts and uses for which the author constructed the text. The manner in which the situation variable is specified is therefore about supposed audience and purpose, and is not simply based on the place where the reading activity is carried out. Many texts used in classrooms are not specifically designed for classroom use. For example, a piece of literary text may typically be read by a 15-year-old in a mother-tongue language or literature class, yet the text was written (presumably) for readers' personal enjoyment and appreciation. Given its original purpose, such a text is classified as *personal* in PISA. As Hubbard (1989) has shown, some kinds of reading usually associated with out-of-school settings for children, such as rules for clubs and records of games, often take place unofficially at school as well. These texts are classified as *public* in PISA. Conversely, textbooks are read both in schools and in homes, and the process and purpose probably differ little from one setting to another. Such texts are classified as *educational* in PISA.

It should be noted that the four categories overlap. In practice, for example, a text may be intended both to delight and to instruct (personal and educational); or to provide professional advice that is also general information (occupational and public). While content is not a variable that is specifically manipulated in this study, by sampling texts across a variety of situations the intent is to maximise the diversity of content that will be included in the PISA reading literacy survey.

Table 2.1 shows the approximate distribution of score points by situation for print and digital reading tasks that will not be finalised until analysis of the main survey data is completed.

Table 2.1
Approximate distribution of score points in reading, by situation

Situation	Percentage of total score points PISA 2012	
	Print	Digital
Personal	36	35
Educational	33	15
Occupational	20	0
Public	11	50
Total	**100**	**100**

Text

Reading requires material for the reader to read. In an assessment, that material – a text (or a set of texts) related to a particular task – must be coherent within itself. That is, the text must be able to stand alone without requiring additional material to make sense to the proficient reader.[2] While it is obvious that there are many different kinds of texts and that any assessment should include a broad range, it is not so obvious that there is an ideal categorisation of kinds of texts. The addition of digital reading to the framework has made this issue still more complex. Since 2009, there have been four main text classifications:

- Medium: print and digital.
- Environment: authored, message-based and mixed.
- Text format: continuous, non-continuous, mixed and multiple.
- Text type: description, narration, exposition, argumentation, instruction and transaction.

The classification of medium – print and digital – is applied to each text as the broadest distinction. Below that classification, the text format and text type categories are applied to all texts, whether print or digital. The environment classification, on the other hand, is only applicable to digital texts.

Medium

Since PISA 2009, an important major categorisation of texts is the classification by medium: print or digital.

Print text usually appears on paper in forms such as single sheets, brochures, magazines and books. The physical status of the printed text encourages (though it does not compel) the reader to approach the content of the text in a particular

sequence. In essence, printed texts have a fixed or static existence. Moreover, in real life and in the assessment context, the extent or amount of the text is immediately visible to the reader.

Digital text may be defined as the display of text through Liquid Crystal Display (LCD), plasma, Thin Film Transistor (TFT) and other electronic devices. For the purposes of PISA, however, digital text is synonymous with *hypertext*: a text or texts with navigation tools and features that make possible and indeed even require non-sequential reading. Each reader constructs a "customised" text from the information encountered at the links he or she follows. In essence, such digital texts have an unfixed, dynamic existence. In the digital medium, typically only a fraction of the available text can be seen at any one time, and often the extent of text available is unknown.

Navigation tools and features help readers to negotiate their way into, around and across texts, through different types of devices: navigation icons, scroll bars, tabs, menus, embedded hyperlinks, text search functions such as Find or Search, and global content representation devices, such as site maps. Navigation features also exist in the print medium (they include tables of contents, indexes, chapter and section headings, headers and footers, page numbers and footnotes) but they play a particularly important role in the digital medium, for at least two reasons. First, due to the reduced display size, digital texts come with devices that let the reader move the reading window over the text page (e.g. scroll bars, buttons, index). Second, typical digital reading activities involve the use of multiple texts, sometimes selecting from a virtually infinite pool. Readers must be familiar with the use of retrieval, indexing and navigation tools for linking between texts.

In the PISA assessment of digital reading, a set of navigation tools and structures has been identified for systematic inclusion in the instruments, as one important component in measuring proficiency in digital reading. This set includes scroll bars, tabs for different websites, lists of hyperlinks[3] displayed in a row, in a column or as a drop-down menu, and embedded text.

Tasks are more or less easy depending on the number of navigation tools that is required to be used, the number of operations or steps required, and the type of tools used. Generally, the larger the number of operations, and the more complex the tool type, the greater the item difficulty. The familiarity, transparency or prominence of navigation tools and features also affects difficulty. Some digital reading tasks require little or even no navigation.

Environment

The *environment* classification applies only to digital texts, and in the PISA reading framework, only computer-based environments are considered. Two broad kinds of digital environment have been identified for assessing the reading of digital texts: *authored* and *message-based* environments. The distinction between them is based on whether or not the reader has the potential to influence the content of the site.

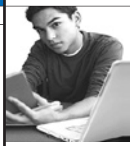
An *authored* environment is one in which the reader is primarily receptive: the content cannot be modified. Readers use these sites mainly for obtaining information. The different types of text within an *authored* environment include home pages, sites publicising events or goods, government information sites, educational sites containing information for students, news sites and online library catalogues.

A *message-based* environment is one in which the reader has the opportunity to add to or change the content. Readers use these sites not only for obtaining information, but also as a way of communicating. Text within a *message-based* environment include e-mail, blogs, chat rooms, web forums and reviews, and online forms.

In practice, as with many of the variables in the reading framework, the environment classifications are not strictly partitioned. Occasionally a task may require integrated use of both authored and message-based texts. Such tasks are classified as *mixed*. Table 2.2 shows the approximate proportion of score points in each environment category.

Table 2.2
Approximate distribution of digital score points in reading, by environment

Environment	Percentage score points in digital reading assessment
Authored	65
Message-based	27
Mixed	8
Total	**100**



Text format

An important classification of texts is the distinction between continuous and non-continuous texts.

Texts in *continuous* and *non-continuous* format appear in both the print and digital media. *Mixed* and *multiple* format texts are also prevalent in both media, particularly so in the digital medium. Each of these four formats is elaborated as follow:

Continuous texts are formed by sentences organised into paragraphs. These may fit into even larger structures, such as sections, chapters, and books (e.g. newspaper reports, essays, novels, short stories, reviews and letters for the print medium, and reviews, blogs and reports in prose for the digital).

Non-continuous texts are organised differently to *continuous* texts, and therefore require a different kind of reading approach. *Non-continuous* texts are most frequently organised in matrix format, composed of a number of lists (Kirsch and Mosenthal, 1990) (e.g. lists, tables, graphs, diagrams, advertisements, schedules, catalogues, indexes and forms).

Many texts in both print and digital media are single, coherent artefacts consisting of a set of elements in both a *continuous* and *non-continuous* format. In well-constructed *mixed* texts, the consituents (e.g. a prose explanation, along with a graph or table) are mutually supportive through coherence and cohesion links at the local and global level. *Mixed* text in the print medium is a common format in magazines, reference books and reports. In the digital medium, authored web pages are typically mixed texts, with combinations of lists, paragraphs of prose, and often graphics. Message-based texts such as online forms, e-mail messages and forums also combine texts that are *continuous* and *non-continuous* in format.

Multiple texts are defined as those that have been generated independently, and make sense independently; they are juxtaposed for a particular occasion or may be loosely linked together for the purposes of the assessment. The relationship between the texts may not be obvious; they may be complementary or may contradict one another. For example, a set of websites from different companies providing travel advice may or may not provide similar directions to tourists. Multiple texts may have a single "pure" format (for example, continuous), or may include both continuous and non-continuous texts.

Table 2.3
Approximate distribution of digital score points in reading, by text format

Text format	Percentage of total score points PISA 2012	
	Print	Digital
Continuous	58	4
Non-continuous	31	11*
Mixed	9	4
Multiple	2	81
Total	**100**	**100**

* Rounded up, the figure is 12% (11.54) but this would make the total 101%. "Approximate" in the title covers this.

Text type

A different categorisation of text is by text type: description, narration, exposition, argumentation, instruction and transaction.

Texts as they are found in the world typically resist categorisation; they are usually not written with rules in mind, and tend to cut across categories. That notwithstanding, in order to ensure that the reading instrument samples across a range of texts that represent different types of reading PISA categorises texts based on their predominant characteristics .

The following classification of texts used in PISA is adapted from the work of Werlich (1976).

Description is the type of text where the information refers to properties of objects in space. The typical questions that descriptive texts provide an answer to are *what* questions (e.g. a depiction of a particular place in a travelogue or diary, a catalogue, a geographical map, an online flight schedule or a description of a feature, function or process in a technical manual).

Narration is the type of text where the information refers to properties of objects in time. Narration typically answers questions relating to *when*, or *in what sequence*. Why characters in stories behave as they do is another important question that narration typically answers (e.g. a novel, a short story, a play, a biography, a comic strip, fictional texts and

a newspaper report of an event). The proportion of narrative texts in the print medium in PISA 2012 is a little greater than that in the previous PISA cycles (2000-09), at about 20% (formerly about 15%).

Exposition is the type of text in which the information is presented as composite concepts or mental constructs, or those elements into which concepts or mental constructs can be analysed. The text provides an explanation of how the different elements interrelate in a meaningful whole and often answers questions about *how* (e.g. a scholarly essay, a diagram showing a model of memory, a graph of population trends, a concept map and an entry in an online encyclopaedia).

Argumentation is the type of text that presents the relationship among concepts or propositions. Argument texts often answer *why* questions. An important sub-classification of argument texts is persuasive and opinionative texts, referring to opinions and points of view. Examples of text in the text type category *argumentation* are a letter to the editor, a poster advertisement, the posts in an online forum and a web-based review of a book or film.

Instruction is the type of text that provides directions on what to do. The text presents directions for certain behaviours in order to complete a task (e.g. a recipe, a series of diagrams showing a procedure for giving first aid, and guidelines for operating digital software).

Transaction represents the kind of text that aims to achieve a specific purpose outlined in the text, such as requesting that something is done, organising a meeting or making a social engagement with a friend. Before the spread of digital communication, this kind of text was a significant component of some kinds of letters and, as an oral exchange, the principal purpose of many phone calls. This text type was not included in Werlich's (1976) categorisation. It was used for the first time in the PISA 2009 framework because of its prevalence in the digital medium (e.g. everyday e-mail and text message exchanges between colleagues or friends that request and confirm arrangements).

Aspect

Whereas navigation tools and features are the visible or physical features that allow readers to negotiate their way into, around and between texts, *aspects* are the mental strategies, approaches or purposes that readers use to negotiate their way into, around and between texts.

Five aspects guide the development of the reading literacy assessment tasks:

- *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.*

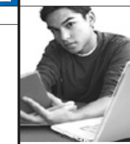
As it is not possible to include sufficient items in the PISA assessment to report on each of the five aspects as a separate subscale, for reporting on reading literacy these five aspects are organised into three broad aspect categories:

- *access and retrieve;*
- *integrate and interpret*; and
- *reflect and evaluate.*

Retrieving information tasks, which focus the reader on separate pieces of information within the text, are assigned to the *access and retrieve* scale.

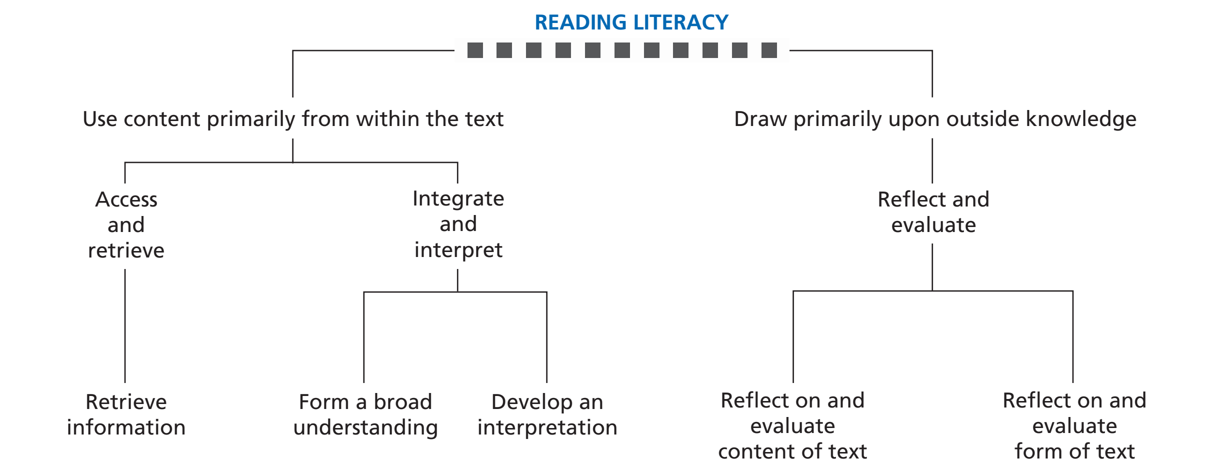
Forming a broad understanding and *developing an interpretation* tasks focus the reader on relationships within a text. Tasks that focus on the whole text require readers to form a broad understanding; tasks that focus on relationships between parts of the text require developing an interpretation. The two are grouped together under *integrate and interpret*.

Tasks addressing the last two aspects, *reflecting on and evaluating the content of a text* and *reflecting on and evaluating the form of a text*, are grouped together into a single *reflect and evaluate* aspect category. Both require the reader to draw primarily on knowledge outside the text and relate it to what is being read. *Reflecting on and evaluating content* tasks are concerned with the notional substance of a text; *reflecting on and evaluating form* tasks are concerned with its structure or formal features.



■ Figure 2.1 ■

Relationship between the reading framework and the aspect subscales



An elaboration of the three broad aspect categories, encompassing tasks in both print and digital media, is given below.

Access and retrieve

Accessing and retrieving involves going to the information space provided and navigating in that space to locate and retrieve one or more distinct pieces of information. Access and retrieve tasks can range from locating the details required by an employer from a job advertisement, to finding a telephone number with several prefix codes, to finding a particular fact to support or disprove a claim someone has made.

While *retrieving* describes the process of selecting the required information, *accessing* describes the process of getting to the place, the information space, where the required information is located. Some items may require retrieving information only, especially in the print medium where the information is immediately visible and where the reader only has to select what is appropriate in a clearly specified information space. On the other hand, some items in the digital medium require little more than accessing (for example, clicking to select an item in a list of search results). However, both processes are involved in most *access and retrieve* tasks in PISA. Difficulty will be determined by several factors including the number of paragraphs, pages or links that need to be used, the amount of information to be processed on any given place, and the specificity and explicitness of the task directions.

Integrate and interpret

Integrating and interpreting involves processing what is read to make internal sense of a text.

Integrating focuses on demonstrating an understanding of the coherence of the text. *Integrating* involves connecting various pieces of information to make meaning, whether it be identifying similarities and differences, making comparisons of degree, or understanding cause and effect relationships.

Interpreting refers to the process of making meaning from something that is not stated. When interpreting, a reader is identifying the underlying assumptions or implications of part or all of the text.

Both *integrating* and *interpreting* are required to *form a broad understanding*. A reader must consider the text as a whole or in a broad perspective. Students may demonstrate initial understanding by identifying the main topic or message or by identifying the general purpose or use of the text.

Both *integrating* and *interpreting* are also involved in *developing an interpretation*, which requires readers to extend their initial broad impressions so that they develop a deeper, more specific or more complete understanding of what they have read. *Integrating* tasks include identifying and listing supporting evidence, and comparing and contrasting information in which the requirement is to draw together two or more pieces of information from the text. In order to process either explicit or implicit information from one or more sources in such tasks, the reader must often infer an intended relationship or category. *Interpreting* tasks may involve drawing an inference from a local context: for example, interpreting the meaning of a word or phrase that gives a particular nuance to the text. This process of comprehension

is also assessed in tasks that require the student to make inferences about the author's intention, and to identify the evidence used to infer that intention.

The relationship between the processes of integration and interpretation may therefore be seen as intimate and interactive. Integrating involves first inferring a relationship within the text (a kind of interpretation), and then bringing pieces of information together, therefore allowing an interpretation to be made that forms a new integrated whole.

Reflect and evaluate

Reflecting and evaluating involves drawing upon knowledge, ideas or attitudes beyond the text in order to relate the information provided within the text to one's own conceptual and experiential frames of reference.

Reflect items may be thought of as those that require readers to consult their own experience or knowledge to compare, contrast or hypothesise. *Evaluate* items are those that ask readers to make a judgment drawing on standards beyond the text.

Reflecting on and evaluating the content of a text requires the reader to connect information in a text to knowledge from outside sources. Readers must also assess the claims made in the text against their own knowledge of the world. Often readers are asked to articulate and defend their own points of view. To do so, readers must be able to develop an understanding of what is said and intended in a text. They must then test that mental representation against what they know and believe on the basis of either prior information, or information found in other texts. Readers must call on supporting evidence from within the text and contrast it with other sources of information, using both general and specific knowledge as well as the ability to reason abstractly.

Reflecting on and evaluating the form of a text requires readers to stand apart from the text, to consider it objectively and to evaluate its quality and appropriateness. Implicit knowledge of text structure, the style typical of different kinds of texts and register play an important role in these tasks. Evaluating how successful an author is in portraying some characteristic or persuading a reader depends not only on substantive knowledge but also on the ability to detect subtleties in language.

Evaluation in the digital medium may take on a slightly different emphasis. The homogeneity of digital text formats (windows, frames, menus, hyperlinks) tends to blur the distinctions across text types. These new features of digital text increase the need for the reader to be aware of authorship, accuracy, quality and credibility of information. As people have access to a broadening universe of information in networked environments, evaluation takes on an increasingly critical role.

To some extent every critical judgment requires the reader to consult his or her own experience; some kinds of reflection, on the other hand, do not require evaluation (for example, comparing personal experience with something described in a text). Thus evaluation might be seen as a subset of reflection.

The aspects of reading in print and digital media

The three broad aspects defined for PISA reading literacy are not conceived of as entirely separate and independent, but rather as interrelated and interdependent. Indeed from a cognitive processing perspective they can be considered semi-hierarchical: it is not possible to interpret or integrate information without having first retrieved it, and it is not possible to reflect on or evaluate information without having made some sort of interpretation. In PISA, however, the framework description of reading aspects distinguishes approaches to reading that are demanded for different contexts and purposes; these are then reflected in assessment tasks that emphasise one or other aspect.

Complex digital reading tasks: Simulating the complexity of real-life reading

While the three aspects do not usually operate entirely independently of one another in either print or digital reading tasks, it is possible to construct relatively simple tasks in which there is a clear emphasis on one or the other aspect. In complex tasks, on the other hand, the process is not so well defined. The reader assimilates the task, and then confronts the problem of interpreting, extrapolating from and evaluating the immediately visible text (for example, the home page of a website) to find relevant information. In an authentic complex task in the digital medium, the reader needs to process the visible information immediately and extrapolate from it: making judgments, synthesising and accessing information in an integrated, recursive sequence.

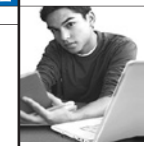


Table 2.4

Approximate distribution of score points in reading, by aspect

Aspect	Percentage of total score points PISA 2012	
	Print	Digital
Access and retrieve	22	19
Integrate and interpret	56	23
Reflect and evaluate	22	19
Complex	0	39*
Total	**100**	**100**

* Rounded (down), the figure is 38% (38.46) but this would make the total 99%. "Approximate" in the title covers this.

Summary of the relationship between printed and digital reading texts and tasks

Table 2.5 presents some of the essential similarities and differences between print and digital reading. One purpose of the table is to describe intrinsic similarities and differences between print reading and digital reading. In many cases the entries under "Print reading" and "Digital reading" are identical. In other places, the descriptions highlight some essential differences in reading between the two media.

A second purpose of the table is to illustrate similarities and differences in what PISA *assesses* in the two media. In some cases it is a matter of prominence and emphasis: square brackets signify that a feature is given relatively little emphasis in the PISA assessment. In other cases the difference is more absolute. While some features exist in both media, they cannot be or are not assessed in PISA. These are printed in blue.

One of the principles in constructing the PISA frameworks and the assessment tasks that operationalise them is to represent the domains authentically. There is no set way of doing this, and in a sense the decisions and selections made are arbitrary, though based on the best judgment of international reading experts. How the domain is described and operationalised, in this and other respects, is determined by a combination of conceptual, empirical and political considerations. The aim in the scoping of the domain outlined above is to explain the basis for building an assessment since PISA 2009 that captures the essence of reading literacy. Such an assessment will in turn yield an array of data from which to report 15-year-olds' reading proficiency in ways that are comprehensive, meaningful and relevant.

ASSESSING READING LITERACY

The previous section outlined the conceptual framework for reading literacy. The concepts in the framework must in turn be represented in tasks and questions in order to collect evidence of students' proficiency in reading literacy.

Building tasks in the print medium

The distribution of tasks across the major framework variables of situation, text and aspect was discussed in the previous section. In this section some of the other major issues in constructing and operationalising the assessment are considered: factors affecting item difficulty, and how difficulty can be manipulated; the choice of response formats; and some issues around coding and scoring.

Factors affecting item difficulty

The difficulty of any reading literacy task depends on an interaction among several variables. Drawing on Kirsch and Mosenthal's work (see for example Kirsch, 2001; Kirsch and Mosenthal, 1990), we can manipulate the difficulty of items by applying knowledge of the following aspect and text format variables.

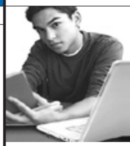
In *access and retrieve* tasks, difficulty is conditioned by the number of pieces of information that the reader needs to locate, by the amount of inference required, by the amount and prominence of competing information, and by the length and complexity of the text.

In *integrate and interpret* tasks, difficulty is affected by the type of interpretation required (for example, making a comparison is easier than finding a contrast); by the number of pieces of information to be considered; by the degree and prominence of competing information in the text; and by the nature of the text: the less familiar and the more abstract the content and the longer and more complex the text, the more difficult the task is likely to be.

Table 2.5
Similarities and differences between print and digital reading, by main framework characteristics

	Print reading	Digital reading
Situations	Personal Public Occupational Educational	Personal Public Occupational Educational
Texts: Environments	Not applicable	Authored Message-based Mixed
Texts: Formats	Continuous Non-continuous [Mixed] [Multiple]	[Continuous] [Non-continuous] [Mixed] Multiple
Texts: Text Type	Argumentation Description Exposition Narration Instruction Transaction	Argumentation Description Exposition Narration Instruction Transaction
Aspects (1)	Access and retrieve Search Orient and navigate in concrete information space e.g. *Go to library, search in a catalogue, find a book* Use navigation tools and structures e.g. *Table of contents; page numbers; glossary* Select and sequence information - low reader control - one sequence of linear reading	Access and retrieve Search Orient and navigate in abstract information space e.g. *Enter URL; user search engines* Use navigation tools and structures e.g. *Menus; embedded hyperlinks* Select and sequence information - high reader control - multiple sequences of linear reading
Aspects (2)	Integrate and interpret Integrate at a lower level of demand: larger portions of text are simultaneously visible (one or two pages) Develop an interpretation Form a broad understanding	Integrate and interpret Integrate at a higher level of demand: limited parts of text are simultaneously visible (limited by screen size) Develop an interpretation Form a broad understanding
Aspects (3)	Reflect and evaluate Pre-evaluate information e.g. *Use table of contents; skim passages, checking for credibility and usefulness* [Evaluate credibility of source - *usually less important due to filtering and preselection in the publishing process*] Evaluate plausibility of content Evaluate coherence and consistency Hypothesise Reflect in relation to personal experience	Reflect and evaluate Pre-evaluate information e.g. *Use menus; skim web pages, checking for credibility and usefulness* Evaluate credibility of source - *usually more important due to lack of filtering and preselection in open environment* Evaluate plausibility of content Evaluate coherence and consistency Hypothesise Reflect in relation to personal experience
Aspects (4)	Complex The range of sources to be consulted is relatively undefined The sequence of steps within the task is undirected e.g. *finding, evaluating and integrating information from multiple printed texts*	Complex The range of sources to be consulted is relatively undefined The sequence of steps within the task is undirected e.g. *finding, evaluating and integrating information from multiple digital texts*

In *reflect and evaluate* tasks, difficulty is affected by the type of reflection or evaluation required (from least to most difficult, the types of reflection are: connecting; explaining and comparing; hypothesising and evaluating); by the nature of the knowledge that the reader needs to bring to the text (a task is more difficult if the reader needs to draw on narrow, specialised knowledge rather than broad and common knowledge); by the relative abstraction and length of the text; and by the depth of understanding of the text required to complete the task.



In tasks relating to *continuous texts*, difficulty is influenced by the length of the text, the explicitness and transparency of its structure, how clearly the parts are related to the general theme, and whether there are text features, such as paragraphs or headings, and discourse markers, such as sequencing words.

In tasks relating to *non-continuous texts*, difficulty is influenced by the amount of information in the text; the list structure (simple lists are easier to negotiate than more complex lists); whether the components are ordered and explicitly organised, for example with labels or special formatting; and whether the information required is in the body of the text or in a separate part, such as a footnote.

Response formats

Coding requirements are shown in Table 2.6 for print score points in relation to the three aspect of reading literacy and for digital score points in relation to the four aspects. Items that require expert judgment consist of open-constructed and short-constructed responses that require expert coding. Items that do not require coder judgment consist of multiple-choice, complex multiple-choice and closed-constructed response items. The closed-constructed response items are those that require the student to generate a response, but require minimal judgment on the part of a coder.

The distribution of item types in print reading does not vary much from one cycle/administration to the next. However, the selection for 2012 has a slightly higher proportion of items that do not require expert coding than in previous cycles: 58% non-expert coded and 42% expert coded in 2012 (compared with 55% and 45% respectively in previous administrations). The same ratio applies to print and to digital reading in PISA 2012.

Table 2.6

Approximate distribution of score points in reading, by coding requirement for each reading aspect

	Print reading			Digital reading		
Aspect	**Expert judgement required**	**No expert judgment required**	**Total**	**Expert judgement required**	**No expert judgment required**	**Total**
Access and retrieve	4	18	22	0	19	19
Integrate and interpret	20	36	56	0	23	23
Reflect and evaluate	18	4	22	15	4	19
Complex	0	0	0	27	12	38
Total	**42**	**58**	**100**	**42**	**58**	**100**

Coding and scoring

Codes are applied to test items, either by a more or less automated process of capturing the alternative chosen by the student for a multiple-choice answer, or by a human judge (expert coder) selecting a code that best captures the kind of response given by a student to an item that requires a constructed response. The code is then converted to a score for the item. For multiple-choice or closed-response format items, the student has either chosen the designated correct answer or not, so the item is scored as 1 (full credit) or 0 (no credit) respectively. For more complex scoring of constructed response items, some answers, even though incomplete, indicate a higher level of reading literacy than inaccurate or incorrect answers, and receive partial credit.

Building tasks in the digital medium

This section considers some of the major issues in constructing and operationalising the digital reading literacy assessment: the relationship between navigation and text processing; analysis of tasks with a view to controlling for item difficulty; response formats; and some issues around coding and scoring. The section ends with a note on the way students' progress through the digital reading assessment is controlled.

Relationship between navigation and text processing in the digital reading assessment

Knowledge of some techniques of navigation and some navigation features are part of being literate in the digital medium. Such skills and knowledge should be regarded as ICT skills that are in conjunction with reading literacy. Both the reading of text, as it is conventionally understood, and the ability to navigate within the digital medium are conceived of as integral to proficiency in digital reading. Each digital reading task includes mental processing devoted to navigation decisions, and textual processing, with more or less weight on each element.

Analysis of digital reading tasks

In order to capture the complexity of the steps that the reader needs to perform in order to arrive at an explicitly called-for response, test developers used a system of analysis to describe the text processing and navigation components of each task.

For any task with a moderate degree of complexity in the digital medium, the reader is likely to have several possible ways of proceeding. For the purposes of describing and analysing subtasks, the test developers imagined an optimally efficient, but comprehensive, sequence of steps, where each step was marked by an *action* (a click on a specified link, a text response in the browser area, a selection from a set of alternatives, or simply scrolling).

For each subtask completed with an action, the following variables were tabulated: text complexity; navigation tool/text used; aspect and description; and action.

Illustrative PISA digital reading items

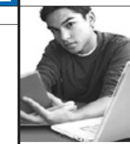
LET'S SPEAK

SCREEN 1A

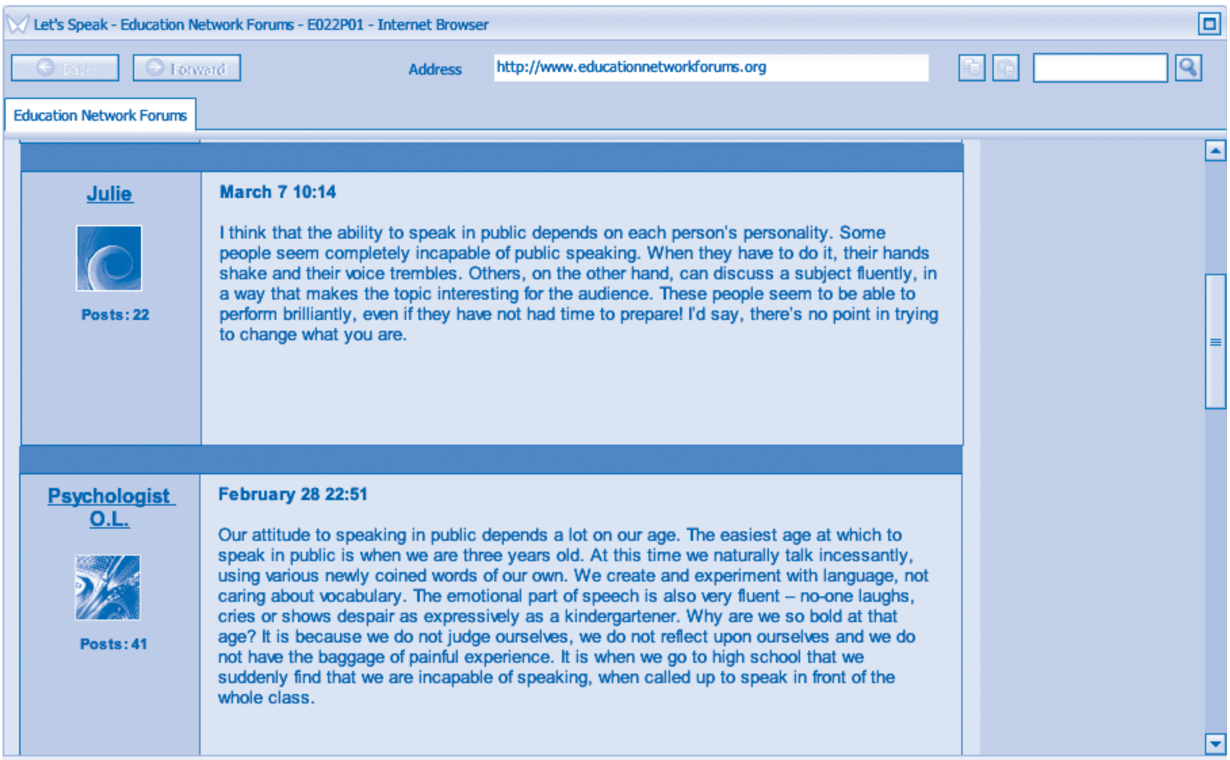
This unit was based on an online discussion forum on the subject of the challenges of speaking in public. The discussion is initiated by Mischa, whose blog entry at the bottom of the discussion forum screen (shown in screen 1E) refers to her terror of speaking in public, to a classroom audience, and asks for help and advice.

The theme of the discussion, set in an educational situation, is an example of a context that would be familiar to most of the PISA students. In terms of text format and text type, *LET'S SPEAK* is categorised as a multiple text, from a number of authors, and argumentative in rhetorical structure. It presents an interactive situation in which the contributors are responding directly to each other. This is a new, or at least much accelerated kind of exchange that is an increasingly prevalent form of communication. In this kind of multiple text, understanding of each text is partly dependent on following the chain of contributions.

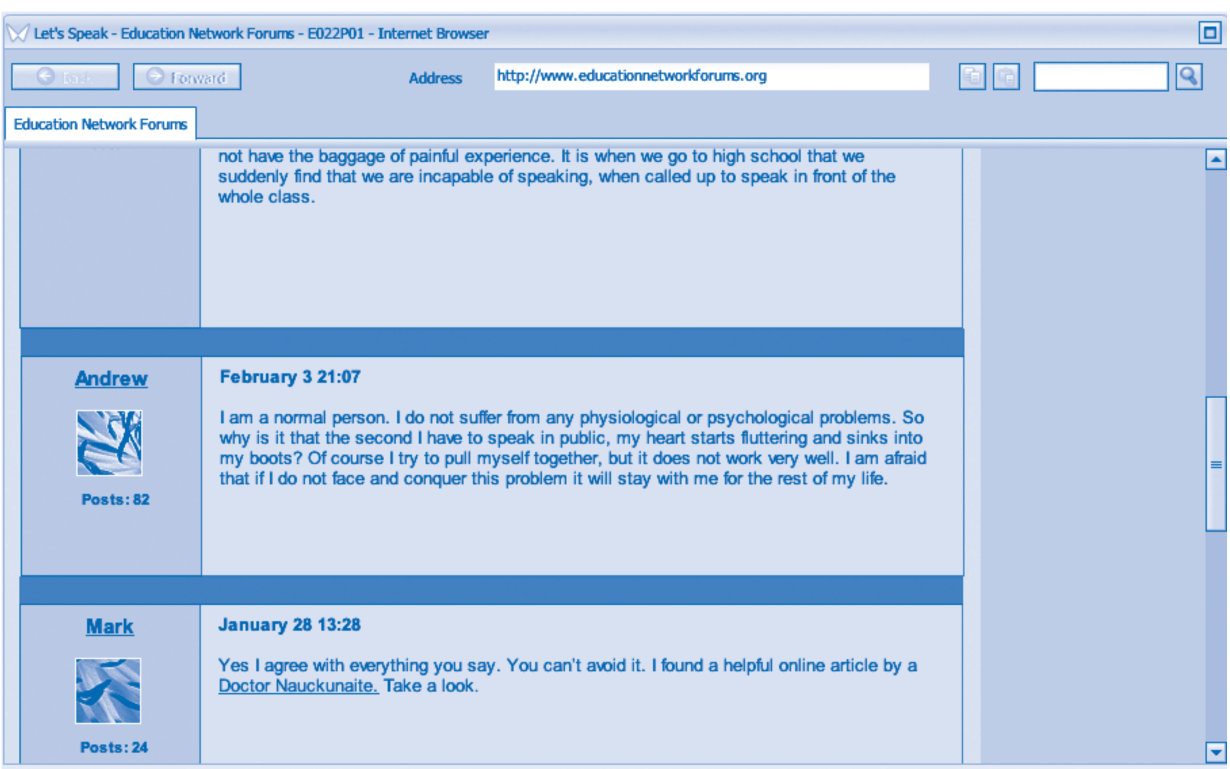
The discussion forum page is quite long, comprised of eight entries. In order to read the initiating entry, it is necessary to scroll down. Screens 1B to 1E show what the reader sees when scrolling down.

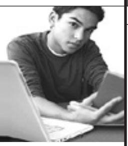


SCREEN 1B

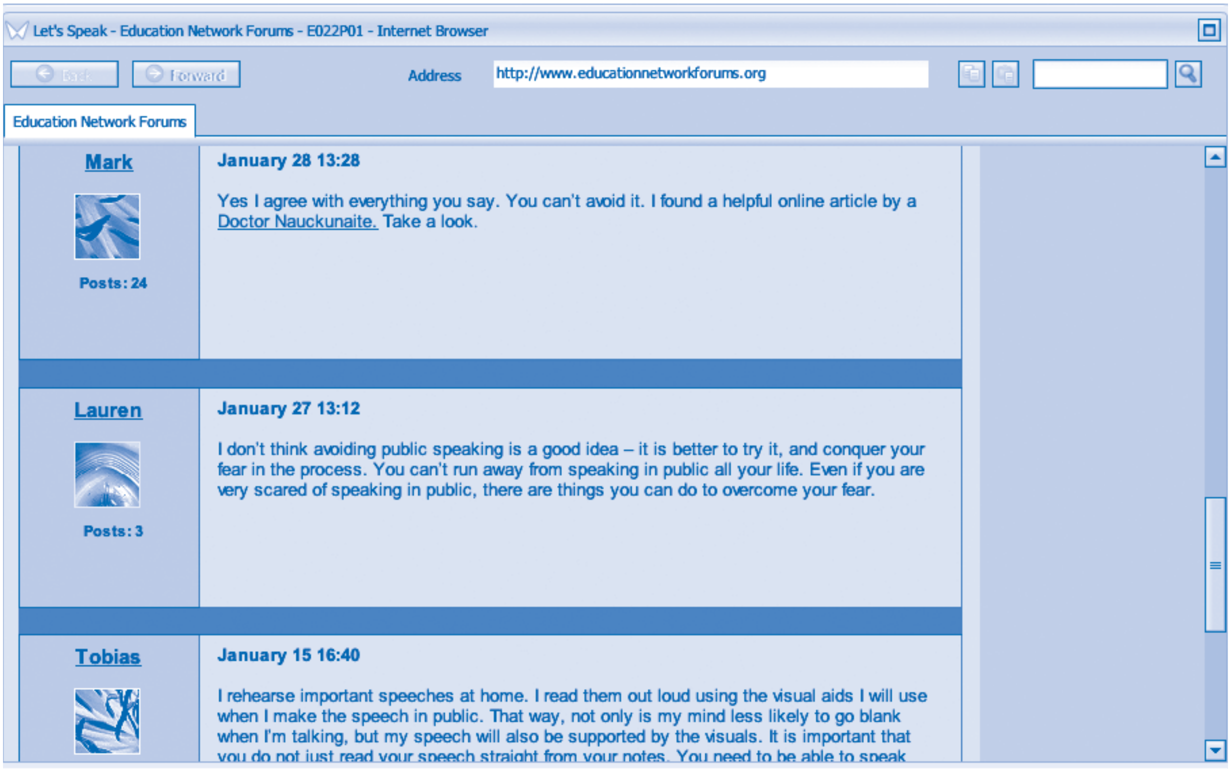


SCREEN 1C

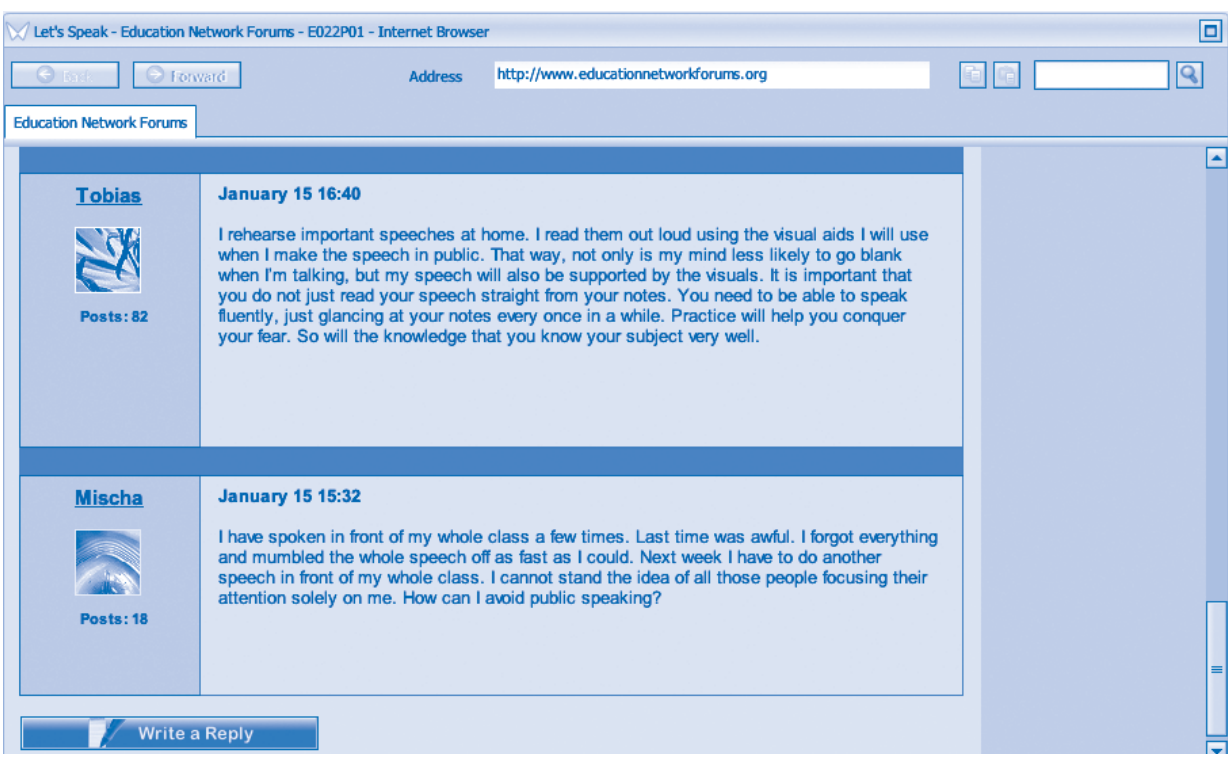




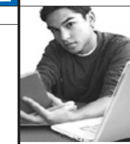
SCREEN 1D



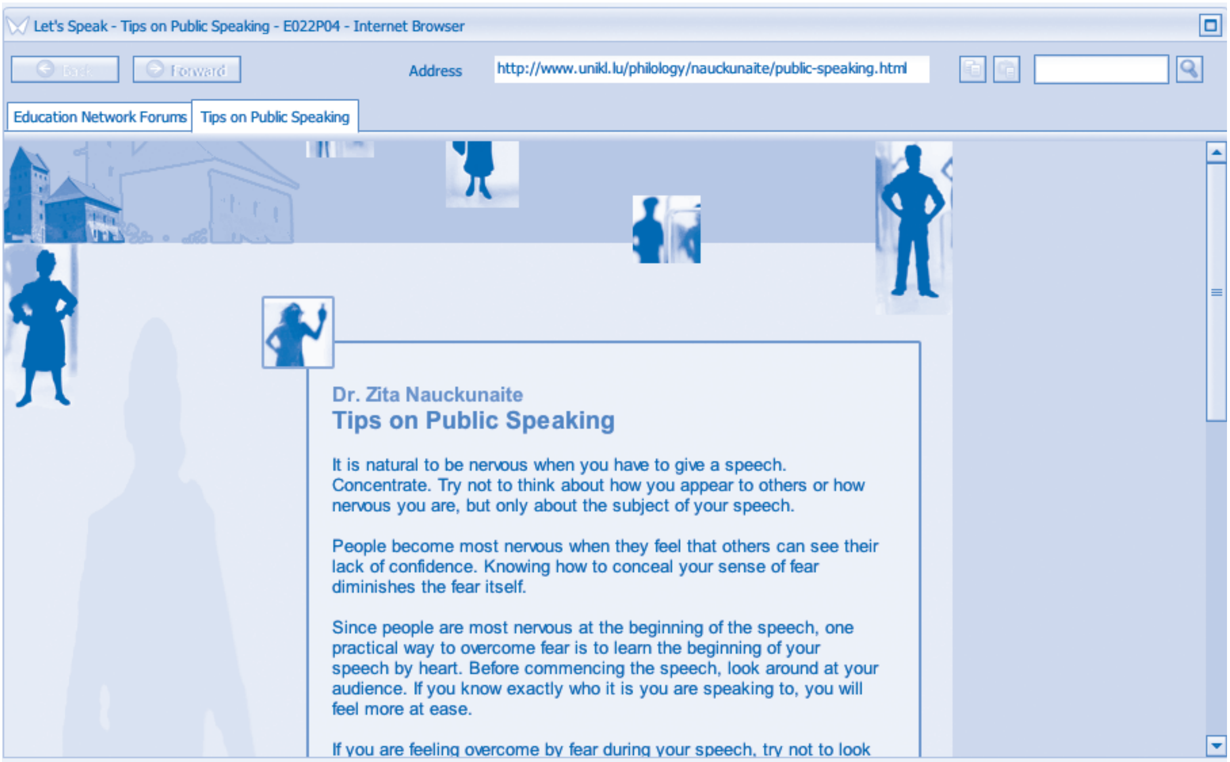
SCREEN 1E



In addition to the starting page, the unit includes only one other piece of stimulus, which is accessed by clicking on an embedded link in one of the blogs that recommends it as "expert advice". The second screen, advice from Doctor Nauckunaite, also requires some scrolling (see screens 2A and 2B).



SCREEN 2A



Dr. Zita Nauckunaite

Tips on Public Speaking

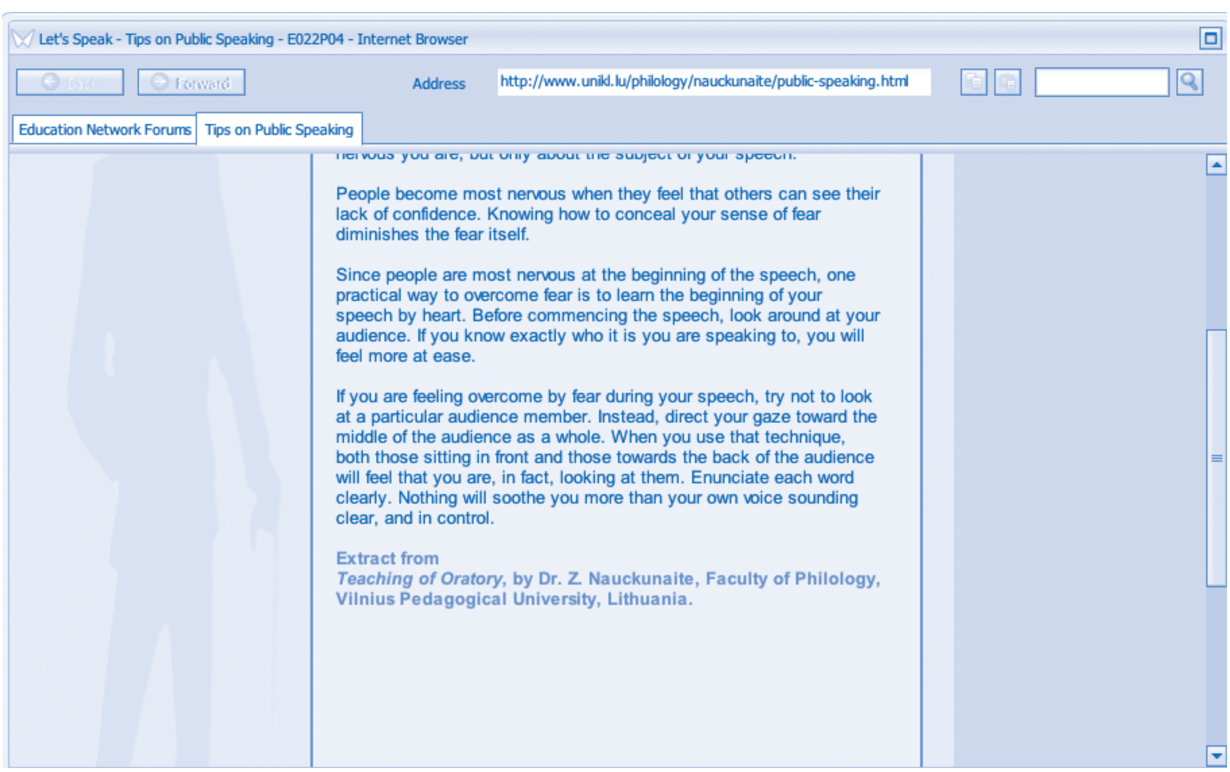
It is natural to be nervous when you have to give a speech. Concentrate. Try not to think about how you appear to others or how nervous you are, but only about the subject of your speech.

People become most nervous when they feel that others can see their lack of confidence. Knowing how to conceal your sense of fear diminishes the fear itself.

Since people are most nervous at the beginning of the speech, one practical way to overcome fear is to learn the beginning of your speech by heart. Before commencing the speech, look around at your audience. If you know exactly who it is you are speaking to, you will feel more at ease.

If you are feeling overcome by fear during your speech, try not to look

SCREEN 2B



nervous you are, but only about the subject of your speech.

People become most nervous when they feel that others can see their lack of confidence. Knowing how to conceal your sense of fear diminishes the fear itself.

Since people are most nervous at the beginning of the speech, one practical way to overcome fear is to learn the beginning of your speech by heart. Before commencing the speech, look around at your audience. If you know exactly who it is you are speaking to, you will feel more at ease.

If you are feeling overcome by fear during your speech, try not to look at a particular audience member. Instead, direct your gaze toward the middle of the audience as a whole. When you use that technique, both those sitting in front and those towards the back of the audience will feel that you are, in fact, looking at them. Enunciate each word clearly. Nothing will soothe you more than your own voice sounding clear, and in control.

Extract from
***Teaching of Oratory*, by Dr. Z. Nauckunaite, Faculty of Philology, Vilnius Pedagogical University, Lithuania.**

This digital reading unit, which was administered in the field trial for PISA 2009, included several tasks that required students to understand the organisation of the website, to identify main ideas both across the blog entries and within an individual entry, and to recognise the existence of conflicting opinions. The final task directed students to read the last entry (at the top of the discussion forum page) in which Mischa has, in an imagined scenario, read all the information provided and is now requesting some final summary advice. This task is reproduced below.

TASK – LET'S SPEAK

Look at Mischa's post for March 10. Click on "Write a Reply" and write a reply to Mischa. In your reply, answer her question about which writer, in your opinion, knows the most about this issue.

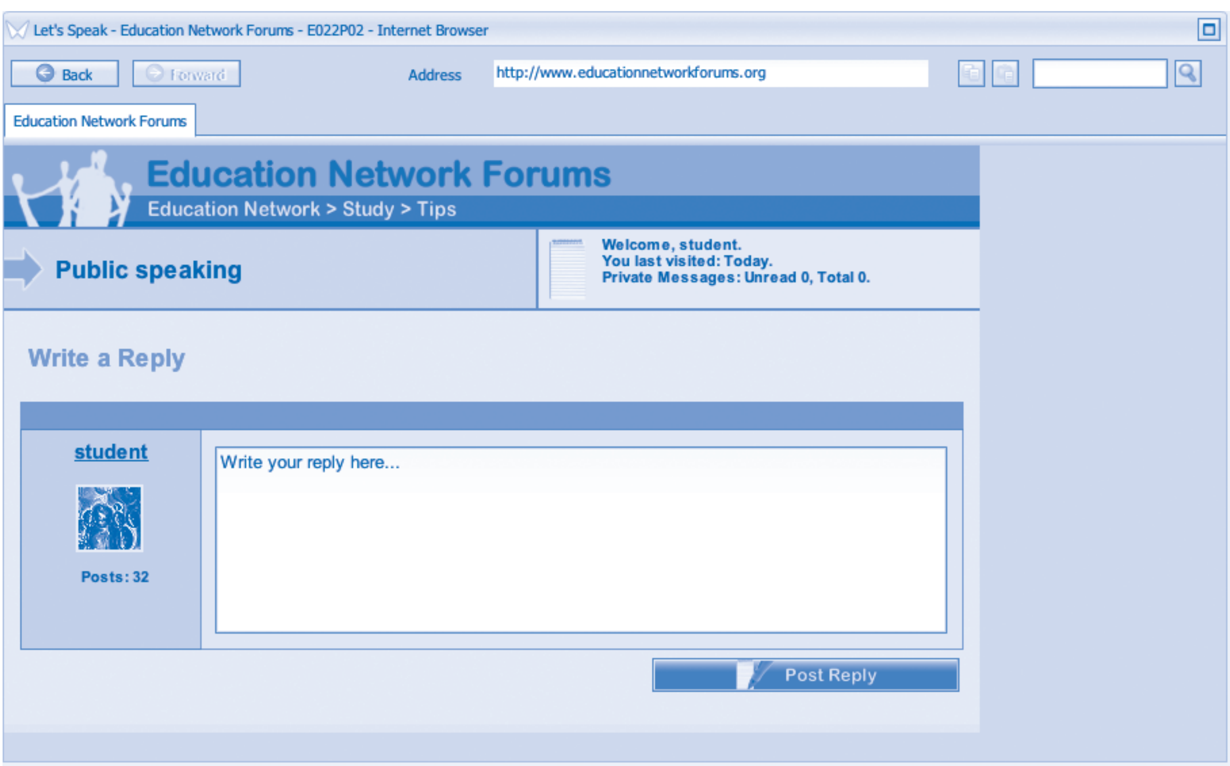
Give a reason for your answer.

Click "Post Reply" to add your reply to the forum.

This is a task that requires access and integration of several pieces of information. Mischa's second blog entry asks the reader to consider and compare four short texts (those of Julie, Tobias, Psych OL and Dr. Nauckunaite). It also requires an evaluation of the contributions, in terms of either their professional credentials, or in terms of the intrinsic quality and persuasiveness of the arguments. It is classified as a *complex* item because it draws significantly on all three aspects: *access and retrieve, integrate and interpret* and *reflect and evaluate*.

An added dimension of the demand of the task is that the student needs to demonstrate some proficiency in handling the formal structure and navigational conventions of the message-based environment by scrolling, clicking on a link that is embedded in the text, and finally clicking on another link (a button) to write a reply. Once the student has clicked on "Write a reply", the screen 3 appears, with an area in which the response can be entered.

SCREEN 3



The coding of this item for the PISA 2009 field trial was based on the text response that the student enters in the "Write a Reply" area. (Note that full credit could be obtained for the response without clicking on "Post Reply" – that detail was added in the interest of authenticity.) However, in developing the item, both the text-processing requirements and the navigational requirements were deliberately manipulated to shape the task for maximum contribution in populating the information space of the assessment. Table 2.7 shows a simplified version of how this *LET'S SPEAK* task can be analysed in terms of its text-processing and navigation components.

Table 2.7

Analysis of a task from the digital reading assessment, *LET'S SPEAK*

Step	Start page / Required text processing / Text complexity rating	Required navigation tools / features	Aspect / text processing description	Action
1	Screen 1A One short argumentative text Rating: medium	Scrollbar	Interpret: form an understanding of the question posed in Misha's message of March 10. Access: infer that the messages of the four entries referred to in Misha's message can be accessed by scrolling, with the first blogger's name ("Julie") already visible.	Scroll down
2	Screen 1B Two short argumentative texts Rating: medium	Scrollbar	Retrieve: match on two names in Mischa's message ("Julie" and "Psychologist OL"). Interpret: form a broad understanding of the main ideas expressed in Julie's and in Psychologist OL's entries. Access: infer that entries of other required bloggers are accessible by scrolling.	Scroll down
3	Screen 1C Two words highlighted in a short argumentative text Rating: low	Embedded link	Access and retrieve: locate Dr Nauckunaite's link embedded in Mark's blog.	Click on embedded link in Mark's blog
4	Screen 2A Formal text comprising expository and instructional elements Rating: medium to high	Scrollbar	Interpret: form a broad understanding of the main ideas expressed in first part of Dr Nauckunaite's page. Access: infer that article continues below bottom of screen.	Scroll down
5	Screen 2B Formal text comprising expository and instructional elements Rating: medium to high	Back button	Interpret: form a broad understanding of the main ideas expressed in second part of Dr Nauckunaite's page. Access: return to discussion forum page using back button (navigation direction provided explicitly in task).	Click on Back button
6	Screens 1A to 1E Eight short argumentative texts (skim) Screen 1E One of two short argumentative texts Rating: medium	Scrollbar	Access: infer that further scrolling is required to locate the last entry named in Mischa's post. Retrieve: match on name in Mischa's message ("Tobias"). Interpret: form a broad understanding of the main idea expressed in Tobias's entry.	Scroll down
7	Screen 1E Write a Reply button Rating: very low	Write a Reply button	Access: access page to write a reply to Mischa	Click on Write a Reply
8	Screen 3 Text box with Write a Reply button [recall of 3 short argumentative texts from screens 1A, 1B and 1C and formal text comprising expository and instructional elements from Screens 2A and 2B] Rating: very high	None	Reflect and evaluate: generate an evaluation of the most authoritative text, combining prior knowledge with information from three short argumentative texts and one longer expository/instructional text.	Text entry response
9 (Optional)	Screen 3 Post Reply button	Post Reply button	Not applicable	Click on Post Reply

For this task, nine distinct steps are described (the last one optional). However, except for step 8, the order of the steps could be changed to achieve exactly the same result. For example, step 1 could be followed by step 3; or the sequence

could begin with step 7 (but by using the "Write a Reply" button shown in screen 1A, and then the "Back" button to return to the main page of the forum). There are many other possible variations in the sequence. As this task illustrates, even with this relatively restricted set of linked pages, readers in the digital medium construct their own text, to a degree, in terms of the order in which they access and process information. The completion of step 8, for full credit, implies good navigation skills in reading digital text (steps 1 to 7), and also strong text-processing skills, since the response requires processing, integration and evaluation of multiple texts, at least one of which is quite demanding (see steps 4 and 5).

Control of the delivery of tasks in the digital reading assessment

As the screen shots for the task from *LET'S SPEAK* show, the interface for a digital reading unit has two distinct areas: a task area in the lower part of the screen, where the question or instruction is located, and a browser area in the upper part of the screen, where the stimulus is located. The task in the task area remains fixed for the duration of an item while the student can navigate around the browser area to access different simulated web pages or applications in the course of completing a task.

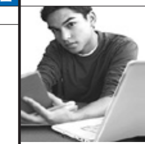
In the digital reading assessment, both units and items within units are delivered in a fixed order, or "lockstep" fashion. The lockstep procedure means that the students are not able to return to an item or unit once they have moved to the next item/unit. A further feature of the task delivery design is that the page that is visible in the browser area at the beginning of each item is fixed: that is, every student sees the same page at the beginning of a given item, regardless of where they finished the previous item. These two features contribute to item independence.

REPORTING PROFICIENCY IN PRINT AND DIGITAL READING

Print reading

PISA reports results in terms of proficiency scales that are interpretable for the purposes of policy. In PISA 2012, reading is a minor domain, and fewer reading items are administered to participating students. A single print reading literacy scale is reported based upon the overall combined scale for print reading.

To capture the progression of complexity and difficulty in PISA 2012, the combined print reading literacy scale is based on the PISA 2009 combined print reading literacy scale and is divided into seven levels. Figure 2.2 describes these seven levels of print reading proficiency. Level 6 is the highest described level of proficiency (Level 5 was the highest level before PISA 2009 reading assessments). The bottom level of measured proficiency is Level 1b (since the PISA 2009 reading assessment, Level 1 was re-labelled as Level 1a and a new level was added, Level 1b, that describes students who would previously have been rated as "below Level 1"). These different levels of proficiency allow countries to know more about the kinds of tasks students with very high and very low reading proficiency are capable of performing. Levels 2, 3, 4 and 5 remain the same in PISA 2012 as in PISA 2000.



■ Figure 2.2 ■

Summary description for the seven levels of proficiency in print reading in PISA 2012

Level	Lower score limit	Percentage of students able to perform tasks at each level or above (OECD average)	Characteristics of tasks
6	698	0.8%	Tasks at this level typically require the reader to make multiple inferences, comparisons and contrasts that are both detailed and precise. They require demonstration of a full and detailed understanding of one or more texts and may involve integrating information from more than one text. Tasks may require the reader to deal with unfamiliar ideas, in the presence of prominent competing information, and to generate abstract categories for interpretations. *Reflect and evaluate* tasks may require the reader to hypothesise about or critically evaluate a complex text on an unfamiliar topic, taking into account multiple criteria or perspectives, and applying sophisticated understandings from beyond the text. A salient condition for *access and retrieve* tasks at this level is precision of analysis and fine attention to detail that is inconspicuous in the texts.
5	626	7.6%	Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of deeply embedded information, inferring which information in the text is relevant. Reflective tasks require critical evaluation or hypothesis, drawing on specialised knowledge. Both interpretative and reflective tasks require a full and detailed understanding of a text whose content or form is unfamiliar. For all aspects of reading, tasks at this level typically involve dealing with concepts that are contrary to expectations.
4	553	28.3%	Tasks at this level that involve retrieving information require the reader to locate and organise several pieces of embedded information. Some tasks at this level require interpreting the meaning of nuances of language in a section of text by taking into account the text as a whole. Other interpretative tasks require understanding and applying categories in an unfamiliar context. Reflective tasks at this level require readers to use formal or public knowledge to hypothesise about or critically evaluate a text. Readers must demonstrate an accurate understanding of long or complex texts whose content or form may be unfamiliar.
3	480	57.2%	Tasks at this level require the reader to locate, and in some cases recognise the relationship between, several pieces of information that must meet multiple conditions. Interpretative tasks at this level require the reader to integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. They need to take into account many features in comparing, contrasting or categorising. Often the required information is not prominent or there is much competing information; or there are other text obstacles, such as ideas that are contrary to expectation or negatively worded. Reflective tasks at this level may require connections, comparisons, and explanations, or they may require the reader to evaluate a feature of the text. Some reflective tasks require readers to demonstrate a fine understanding of the text in relation to familiar, everyday knowledge. Other tasks do not require detailed text comprehension but require the reader to draw on less common knowledge.
2	407	81.2%	Some tasks at this level require the reader to locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. Others require recognising the main idea in a text, understanding relationships, or construing meaning within a limited part of the text when the information is not prominent and the reader must make low level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflective tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.
1a	335	94.3%	Tasks at this level require the reader to locate one or more independent pieces of explicitly stated information; to recognise the main theme or author's purpose in a text about a familiar topic, or to make a simple connection between information in the text and common, everyday knowledge. Typically the required information in the text is prominent and there is little, if any, competing information. The reader is explicitly directed to consider relevant factors in the task and in the text.
1b	262	98.9%	Tasks at this level require the reader to locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. The text typically provides support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. In tasks requiring interpretation the reader may need to make simple connections between adjacent pieces of information.

Digital reading

For those countries that chose to implement the assessment of digital reading, an additional scale, based only on digital reading tasks, was created since PISA 2009 and started a new trend line. Given the relatively small number of items in the pool for PISA 2012 (as for PISA 2009), the range of difficulty of digital reading tasks allows for the description of four levels of reading proficiency: Level 2, Level 3, Level 4 and Level 5 or above. Figure 2.3 describes the four level of proficiency in digital reading. Students with proficiency within the range of Level 2 are likely to be able to successfully complete tasks within that band of difficulty, but are unlikely to be able to complete tasks at higher levels. Students with scores within the range of Level 4 are likely to be able to successfully complete tasks located at that level and at the lower levels.

■ Figure 2.3 ■

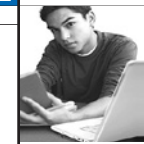
Summary description for the four levels of proficiency in digital reading in PISA 2012

Level	Lower score limit	Percentage of students able to perform tasks at each level or above (OECD average)	Characteristics of tasks
5 or above	626	7.8%	Tasks at this level typically require the reader to locate, analyse and critically evaluate information, related to an unfamiliar context, in the presence of ambiguity. They require generating criteria to evaluate the text. Tasks may require navigation across multiple sites without explicit direction, and detailed interrogation of texts in a variety of formats.
4	553	30.3%	Tasks at this level may require the reader to evaluate information from several sources, navigating across several sites comprising texts in a variety of formats, and generating criteria for evaluation in relation to a familiar, personal or practical context. Other tasks at this level demand that the reader interpret complex information according to well-defined criteria in a scientific or technical context.
3	480	60.7%	Tasks at this level require that the reader integrate information, either by navigating across several sites to find well-defined target information, or by generating simple categories when the task is not explicitly stated. Where evaluation is called for, only the information that is most directly accessible or only part of the available information is required.
2	407	83.1%	Tasks at this level typically require the reader to locate and interpret information that is well-defined, usually relating to familiar contexts. They may require navigation across a limited number of sites and the application of web-based navigation tools such as drop-down menus, where explicit directions are provided or only low-level inference is called for. Tasks may require integrating information presented in different formats, recognising examples that fit clearly defined categories.

SUMMARY

An essential function of PISA is to provide information to policy makers about trends over time. Since PISA 2009, the construction of a scale and subscales that are based entirely on print reading tasks has helped to record and analyse trends. A different set of scales is built to report on the digital reading assessment and, where possible, to report the combined results of print and digital reading assessments, therefore providing the basis for establishing new trend lines for future cycles. In anticipating a range of options for reporting, the PISA reading literacy framework and assessment provide a rich array of data to inform the work of policy makers, educators, and researchers.

The PISA 2012 reading framework has not changed from the PISA 2009 framework. The notion of reading literacy in PISA goes beyond the simple measurement of a student's capacity to decode and understand literal information. Reading literacy in PISA also involves understanding, using, reflecting on and engaging with written texts, both to achieve personal goals and to participate actively in society.



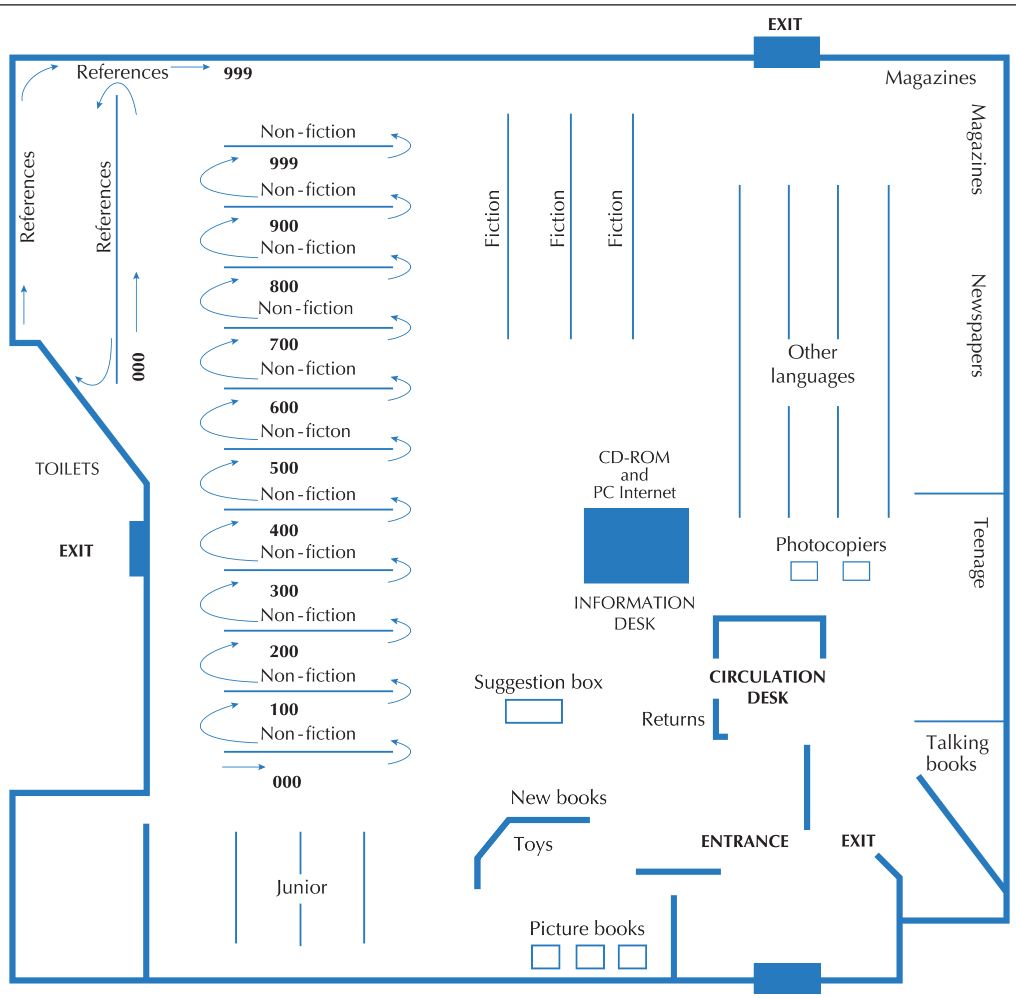
ILLUSTRATIVE PISA PRINT READING ITEMS

LIBRARY MAP

The library map that forms the basis of this unit is an example of a kind of everyday non-continuous text that is often encountered in work, personal, public and educational settings. The context of this example is defined as public because the map relates to the activities of a community (a public library) and assumes anonymous contact with the reader. In terms of text type, the map is classified as description, since the information it contains refers to properties of objects in space and their relationship to one another.

■ Figure 2.4 ■

Items for the unit *LIBRARY MAP*



QUESTION 1

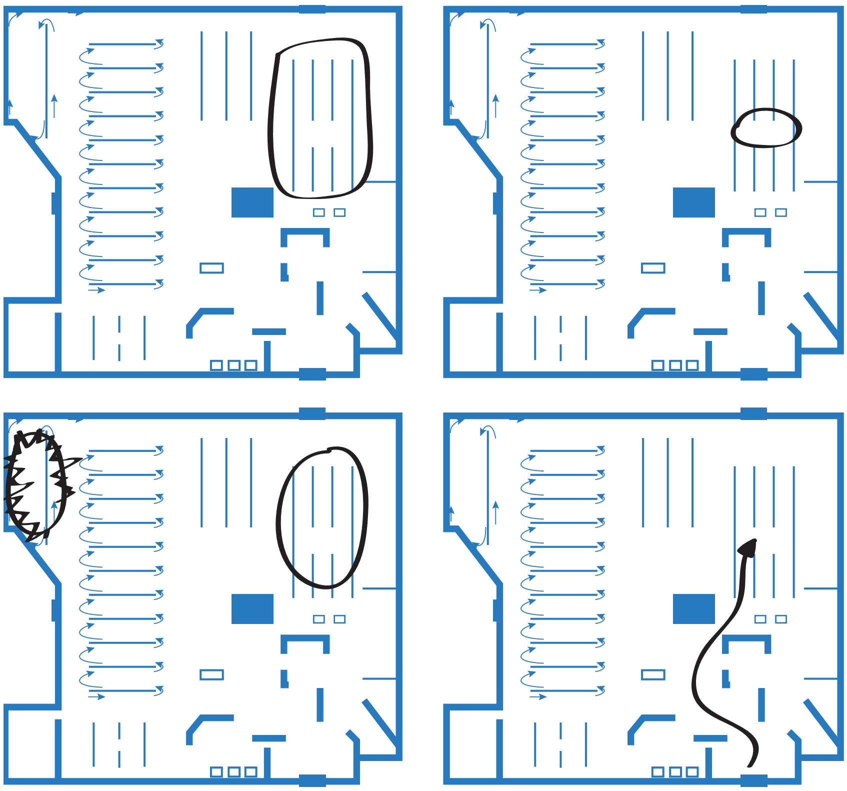
For school you need to read a novel in French.
On the map draw a circle around the section where you would be most likely to find a suitable book to borrow.

The framework characteristics are described below:

- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Description
- Aspect: Access and retrieve: Retrieve information
- Question intent: Locate information that matches on one factor using low-level inference
- Item format: Short response

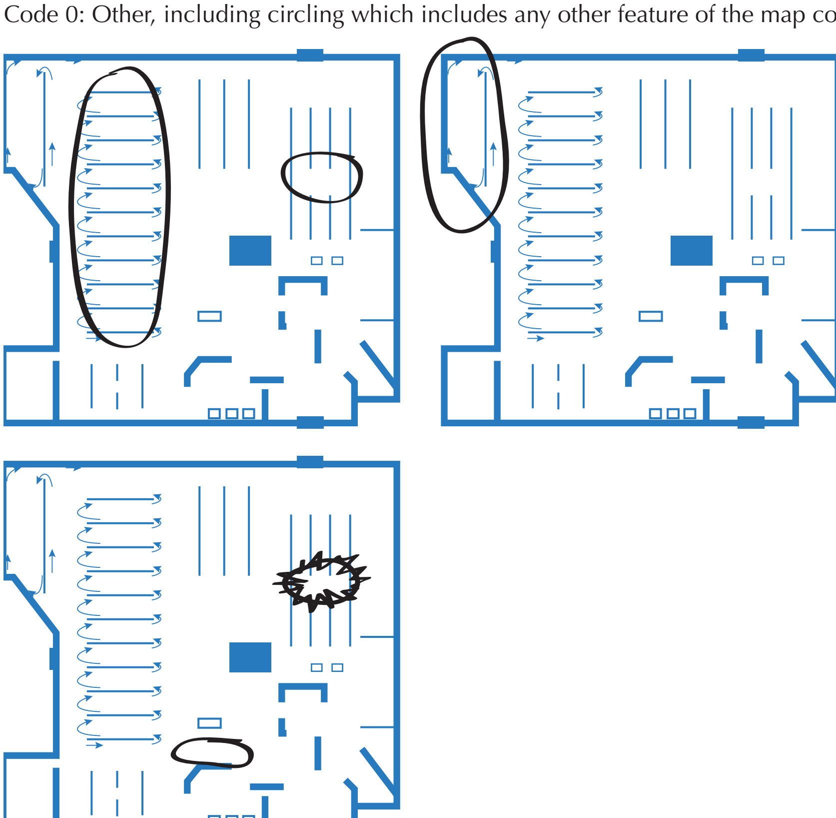
Full credit

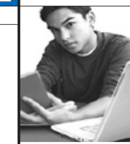
Code 1: Circles the words "other languages" or the lines (shelves) near the words.



No credit

Code 0: Other, including circling which includes any other feature of the map completely.





Code 9: Missing.

This short response item requires that the reader search for, locate and select relevant information from the information space: in this case, a map. The required information is found in a single location rather than multiple locations, a factor that is likely to reduce difficulty. On the other hand, the match between the words in the task and the caption on the map is not literal: the reader must make an inference to categorise "French" as "Other languages". (A translation and adaptation note instructed that in national versions of the item the language referred to in the item should be a foreign language commonly taught in schools.) Nevertheless, this is a rather easy item, with more than four fifths of the students in the field trial able to identify the right section of the library. As indicated in the full credit examples provided with the coding guide, students could mark the text in a number of different ways to show their answer. Although the question specifies that a circle should be drawn to show the answer, the format of the response is not the critical criterion for awarding credit: what is critical is whether or not the response clearly meets the intent of the question – "locating information that matches on one factor using low-level inference".

QUESTION 2A

Where are *New books* located?

A. In the fiction section.

B. In the non-fiction section.

C. Near the entrance.

D. Near the information desk.

The correct answer is C: "Near the entrance". This question is for information only and will not independently contribute to the student's score. The answer is taken into account in assessing the response to Question 2B.

QUESTION 2B

Explain why this location might have been chosen for *New books*.

The framework characteristics are described below:

- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Description
- Aspect: Reflect and evaluate: Reflect on and evaluate the content of a text
- Question intent: Hypothesise about the location of a feature of a map drawing on personal knowledge and experience
- Item format: Open-constructed response

Full credit

Code 2: Answer to Part A correct. Gives an explanation which is consistent with the answer "near the entrance".

- People will see them as soon as they walk in.
- They are away from the other books, and people will find them easily.
- So people can look at them first. [Implies recognition that the new books are near the entrance.]
- So they are very visible.
- They are clearly visible and not hidden away among the bookshelves so that you have to search for them.
- You pass it on your way to fiction.

OR: Answer to previous question Part A correct. Gives an explanation which shows understanding of the location of the new books in relation to a part of the library other than the entrance.

- It gives children a chance to play while adults look around. [Recognises that the new books are near the Toys section.]
- When people are returning books they will see the new ones.

Partial credit

Code 1: Answer to Part A incorrect. Gives an explanation which is consistent with the answer given for previous question.

- [Answer to Part A: In the fiction section.] Because this is the part of the library that most people would be using, so they would notice the new books.
- [Answer to Part A: Near the information desk.] Because they are next to the Information Desk, the librarian can answer questions about them.

No credit

Code 0: Gives insufficient or vague explanation regardless of whether answer to Part A is correct or incorrect.

- Because it's the best place.
- They are near the entrance too. [States where the new books are, without offering explanation.]
- The New books are near the suggestion box. [States where the new books are, without offering explanation.]

OR: Shows inaccurate comprehension of the material or gives an implausible or irrelevant explanation, regardless of whether answer to Part A is correct or incorrect.

- So people would notice them when they were looking at the newspapers. [Inaccurate, implies that new books are near the newspapers.]
- Because there is nowhere else to put them. [Implausible]
- Some people like to read new books. [Answer is irrelevant to question.]
- [Answer to Part A: In the fiction section.] So that they are easy to find. [Answer irrelevant to answer given for Part A]

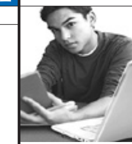
Code 9: Missing.

The coding rules for this task are somewhat complicated. Students are asked two questions – one multiple-choice and one constructed response – but only the second of these is coded directly. As this task contributes to the *reflect and evaluate* scale, the multiple-choice component, which predominantly requires retrieval of information, does not earn any credit on its own. However, the multiple-choice question is taken into account in the coding of the second, constructed response question.

To gain full credit, the response must include both accurate reading of the map (locating the New books near to the entrance) and a hypothesis about the reason for locating the New books in that position. To make such an hypothesis, readers need to consult their own experience or knowledge – in this case about the way libraries work and the way they are used by the public. In the PISA context, the outside knowledge required is intended to be within the expected range of 15-year-olds' experiences.

Students receive only partial credit if they have failed to correctly locate the New books on the map, but have given a plausible hypothesis about the reason for locating New books in a particular position. Like the full credit responses, this kind of response fulfils the intent of reflecting on content that is the main thrust of this task.

This was an easy item, with over four fifths of the students in the field trial gaining full credit.



SUPERMARKET NOTICE

This public notice consists of a very short text that has an everyday function: to warn about the possible danger of a product to consumers and to give advice to return the product for a refund. While the formatting of the stimulus reflects the international standard for product recall notices, many students may not have seen this kind of notice. Nevertheless, the content of the warning is clearly set out and a minimum number of words is used. Lemon biscuits were chosen as the product because of their familiarity and likely appeal. In developing very short easy items, the test developers sought to use simple pieces of stimulus with familiar content. This was not only to make the cognitive load of the items lighter, but also to present texts that were unlikely to intimidate students with low reading proficiency, since such readers can easily be discouraged from even attempting to read something that they believe looks too hard or too long. The text format classification of the supermarket notice is *non-continuous*, as it consists of a list of described features. In terms of text type, the notice is instructional: it provides directions on what to do if you have bought the product.

■ Figure 2.5 ■

Items for the unit *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 1

What is the purpose of this notice?

A. To advertise Lemon Cream Biscuits.

B. To tell people when the biscuits were made.

C. To warn people about the biscuits.

D. To explain where to buy Lemon Cream Biscuits.

The framework characteristics are described below:

- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Instruction
- Aspect: Integrate and interpret: Form a broad understanding
- Question intent : Recognise the main idea of a short text by combining adjacent pieces of information
- Item format: Multiple choice

Full credit

Code 1: C. To warn people about the biscuits.

No credit

Code 0: Other responses.

Code 9: Missing.

To answer this question correctly, students must form a global understanding of the text to recognise its overall purpose. In particular, to reject distractors A and D, students must recognise that although the text is about a particular product, it is not an advertisement, but a warning. This item was easy. The easiness of this item comes in part from the fact that the whole text is very short.

QUESTION 2

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

....

The framework characteristics are described below:

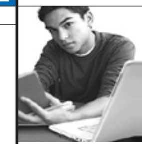
- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Instruction
- Aspect: Access and retrieve: Retrieve information
- Question intent: Locate a synonymous match in a short text
- Item format: Closed-constructed response

Full credit

Code 1: Fine Foods Ltd.

No credit

Code 0: Other responses.



Code 9: Missing.

To answer this question successfully the student needs to locate a single explicitly stated piece of information in the text, using a synonymous match between the task direction and the text (company / manufacturer). The fact that the whole text is very short, and that the needed information is near the beginning of the text, adds to the easiness of the task. The response format for the task is described as closed constructed response, since only one answer (with a small range of variants: Fine Foods or Fine Foods Ltd.) is given full credit.

QUESTION 3

What would **you** do if you had bought these biscuits?

....

Why would you do this?

Use information from the text to support your answer.

....

The framework characteristics are described below:

- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Instruction
- Aspect: Reflect and evaluate: Reflect on and evaluate the content of a text
- Question intent: Hypothesise about a personal course of action in response to the information in a text
- Item format: Open-constructed response

Full credit

Code 1: 3A: Provides a response that is consistent with an understanding that the biscuits may be returned with a refund. May refer to eating the biscuits, not eating the biscuits, returning them or getting rid of them in some other way AND 3B: Gives an explanation consistent with the text and the response in 3A. Must be consistent with the idea that the peanuts pose a potential threat.

- (3A)
  Ask for my money back.
  (3B)
  It tells me to.
  I'm allergic to peanuts.
  They did something wrong.
  There might be something (else) wrong.
  I don't like peanuts.
- (3A)
  Throw them away.
  (3B)
  I'm allergic to peanuts.
  There might be something wrong.
- (3A)
  Eat them.
  (3B)
  Peanuts won't harm me.
  I'm not allergic to peanuts.
  I like peanuts.

- (3A)
  Give them to my classmate,
  (3B)
  She's not allergic to peanuts.
- (3A)
  Nothing.
  (3B)
  I'm not allergic to peanuts.
  I can't be bothered to go back to the shop.

3A: Quotes from or paraphrases an appropriate section of the text without further explanation (implying that the text tells you what to do and that no further explanation is required).
3B: No response.

- (3A) Return the product to the place of purchase for a full refund. Or call 1800 034 241 for further information.
  (3B) (no response)
- (3A) Return the product to the place of purchase for a full refund.
  (3B) (no response)
- (3A) Call 1800 034 241 for further information.
  (3B) (no response)
- (3A) Call the number for more information.
  (3B) (no response)

3A: No response AND 3B: Gives explanation for taking no action. Must be consistent with the idea that the peanuts pose a potential threat.

- (3A) (no response)
  (3B) I'm not allergic to peanuts.
- (3A) (no response)
  (3B) I can't be bothered to go back to the shop.

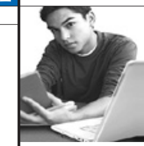
No credit

Code 0: Gives an insufficient or vague response.

- (3A) I don't know
  (3B) they might have peanuts
- (3A) eat them
  (3B) there might be peanuts

Shows inaccurate comprehension of the material or gives an implausible or irrelevant response.

- (3A) (no response)
  (3B) check them for nuts.
- (3A) eat them.
  (3B) they look good enough to eat.
- (3A) give them to someone.
  (3B) it doesn't matter.
- (3A) (no response)
  (3B) I'm allergic to peanuts.
- (3A) (no response)
  (3B) peanuts can be dangerous.
- (3A) throw them away.
  (3B) They're past their Best before date.



Code 9: Missing.

This question requires students to hypothesise about their likely personal response to the information in the text. Since the question requires a judgement based on personal preferences, or likely behaviours, the question is classified as *reflect and evaluate*. The coding guide indicates that a wide range of responses can receive full credit, so long as the response is consistent with two central ideas of the text: firstly, that it is possible to return the biscuits, and secondly that the biscuits pose a potential threat. The item is easy, with over four-fifths of the field trial respondents gaining full credit. The easiness of the item can be explained in part by the low level of reflection to be done: no specialised knowledge is required in order to explain a personal preference about a course of action regarding the familiar topic of food.

QUESTION 4

Why does the notice include "Best before" dates?

....

The framework characteristics are described below:

- Situation: Public
- Medium: Print
- Text format: Non-continuous
- Text type: Instruction
- Aspect: Integrate and interpret: Develop an interpretation
- Question intent : Identify the purpose of a conventional feature included in a short text
- Item format: Open-constructed response

Full credit

Code 1: Refers to the fact that the Best before dates identify the batches of biscuits that are affected.

- to identify the batch(es).
- so you know which packets have peanuts.

No credit

Code 0: Refers to when the biscuits should be eaten.

- because that's when you eat them.
- to tell you when to eat the biscuits.
- so you don't keep them too long.
- to tell you when they expire.

Gives an insufficient or vague response.

- it's the date.

Shows inaccurate comprehension of the material or gives an implausible or irrelevant response.

- so you know when the notice is irrelevant.

Code 9: Missing.

This question was answered correctly by less than one-third of students. Given the shortness and simplicity of the text, this illustrates the fact that the characteristics of a text only partly explain the difficulty of an item. The question requires students to identify the purpose of a specified part of the text, namely, the "Best before" dates. The difficulty of the item comes from the fact that students must focus on the purpose of the feature in this particular text. Students who answer by giving the usual purpose of this feature (that is, to tell the consumer when the product should be used by) do not receive credit for this item. In this respect the full credit response is contrary to expectations, an established marker of item difficulty.

DESTINATION BUENOS AIRES

Destination Buenos Aires is an extract from Antoine de Saint-Exupéry's 1931 novel *Vol de Nuit* (published in English as *Night Flight*). The only addition to the original text for its appearance in PISA was an explanatory footnote relating to "Patagonia", as students would certainly have differing levels of familiarity with this place name. The explanation gives context which might help students to negotiate the text. The extract takes place at a landing-ground in Buenos Aires, and is a self-contained portrait of Rivière, a man weighed down by the responsibility of his job. Though the novel was written in 1931, the human themes remain familiar.

■ Figure 2.6 ■

Items for the unit *DESTINATION BUENOS AIRES*

And so the three mail planes from Patagonia,[1] Chile and Paraguay were returning from the South, the West and the North to Buenos Aires. Their cargo was awaited there so that the plane for Europe could take off, around midnight.

Three pilots, each behind an engine casing heavy as a barge, lost in the night, were contemplating their flight and, approaching the immense city, would descend slowly out of their stormy or calm sky, like strange peasants descending from their mountain.

Rivière, who was responsible for the entire operation, was pacing up and down on the Buenos Aires landing-ground. He remained silent, for until the three planes had arrived, the day held a sense of foreboding for him. Minute by minute, as the telegrams reached him, Rivière was conscious that he was snatching something from fate, gradually reducing the unknown, hauling in his crews out of the night, towards the shore.

One of the men came up to Rivière to give him a radioed message:

Chile mail reports that he can see the lights of Buenos Aires.

Good.

Before long, Rivière would hear this plane; already the night was surrendering one of them, just as a sea, swollen with ebbing and flowing and mysteries, surrenders to the shore the treasure it has tossed around for such a long time. And later on, it would give back the other two.

Then this day's work would be over. Then the worn-out crews would go and sleep, to be replaced by fresh crews. But Rivière would have no rest: the mail from Europe, in its turn, would fill him with apprehension. And so it would always be. Always.

Antoine de Saint-Exupéry, *Vol de Nuit* © Éditions Gallimard

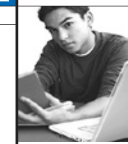
1. Southern region of Chile and Argentina.

QUESTION 1

How does Rivière feel about his job? Use the text to give a reason to support your answer.

The framework characteristics are described below:

- Situation: Personal
- Medium: Print
- Text format: Continuous
- Text type: Narration
- Aspect: Integrate and interpret: Develop an interpretation
- Question intent: Link information across a narrative to generalise about a character's state of mind, providing evidence to support the generalisation
- Item format: Open-constructed response



Full credit

Code 2: Describes Rivière's feeling about his job by referring to stress, persistence, being burdened, or being committed to doing his duty; AND gives an explanation referring to a relevant section of the text. May refer to the text generally, or may paraphrase or quote the text directly. The quotation must match the stated emotion.

- He is overwhelmed by it all, you can see in the last line, he never gets to rest.
- He is stressed. The day has "held a sense of foreboding for him".
- He is weighed down by it. All day he worries about those three planes, then he has to worry about the Europe one!
- He is resigned. You can see from that last "always" that he thinks things will never change.
- He really cares about his job. He can't relax until he knows that everyone is safe. [Includes a general reference to the text.]

Partial credit

Code 1: Describes Rivière's feeling about his job by referring to stress, persistence, being burdened, or being committed to doing his duty, without an explanation that refers to the text.

- He feels really responsible for the things that happen.
- He's stressed.

No credit

Code 0: Gives an insufficient or vague response.

Shows inaccurate comprehension of the material or gives an implausible or irrelevant response.

- He likes his job because he is in control of lots of things. [not supported by the text]
- He thinks it is cool because he can watch planes. [not supported by the text]

Code 9: Missing.

The coding guide for this item shows that there are two kinds of response that receive credit. Full credit responses are those which accurately respond to the question and give an explanation using the text. Partial credit responses are those which accurately respond to the question, but fail to give an explanation for the response. The partial credit code recognises that an incomplete answer is superior to an inaccurate one. In the field trial, less than half of the students received full credit for this item, but an additional one quarter received partial credit, meaning that about three-quarters of students received some credit (either full or partial) for this item. This question is classified as *integrate and interpret*, because although students are required to generate a response that is not given explicitly in the text, all the information necessary to answer the question is contained within the text.

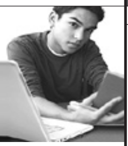
QUESTION 2

"Destination Buenos Aires" was written in 1931. Do you think that nowadays Rivière's concerns would be similar? Give a reason for your answer.

....

The framework characteristics are described below:

- Situation: Personal
- Medium: Print
- Text format: Continuous
- Text type: Narration
- Aspect: Reflect and evaluate: Reflect on and evaluate the content of a text
- Question intent : Hypothesise about the effect on a character of a change in a narrative's context
- Item format: Open-constructed response



Full credit

Code 1: Answers (or implies) Yes OR No and refers to a time-based comparison AND supports their answer. May refer to material concerns such as technological progress or improvements in security OR to psychological concerns such as anxiety. Answer must be consistent with an accurate reading of the text.

- Now, pilots (planes) have very sophisticated tools intended for orientation, making up for technical issue when the weather conditions are bad.
- No, nowadays, planes have radars and automatic piloting systems, which can help them to escape from dangerous situations.
- Yes, planes are still dangerous, just like any other means of transport. The risks of crash or engine failure are never eradicated.
- Now, new technologies and technical progress are very important, in the planes as well as on the ground.
- Yes, there is still a risk of crashing.
- No, before, there was no fear of terrorist attacks.

No credit

Code 0: Gives an insufficient or vague response.

- No, the fears are different today.
- Yes, some progress has been made.
- In a way, yes, but in the modern day context. [vague]
- Over the years, people would have changed it. [vague]

Shows inaccurate comprehension of the material or gives an implausible or irrelevant response.

- No, because you don't travel by night nowadays. [inaccurate about the world]
- No, because nowadays, pilots are much better trained. [irrelevant]
- No, Rivière is really happy with his job but nowadays there are terrorists to worry about. [inaccurate reading of the text]

Code 9: Missing.

This item was moderately difficult. Just over one-half of students answered correctly. The item requires students to reflect on the context in which a text was written and compare that context to their own. The object of the question is to encourage reflection. Therefore, so long as the response is consistent with an accurate reading of the text, and expresses a plausible position about the modern day context, a wide range of responses receive full credit, regardless of the position adopted.

QUESTION 3

What happens to the main character in this text?

A. He has an unpleasant surprise.

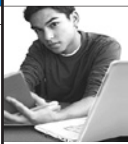
B. He decides to change his job.

C. He waits for something to happen.

D. He learns to listen to others.

The framework characteristics are described below:

- Situation: Personal
- Medium: Print
- Text format: Continuous
- Text type: Narration
- Aspect: Integrate and interpret: Form a broad understanding
- Question intent : Recognise the main action in a narrative text
- Item format: Multiple choice



Full credit

Code 1: C. He waits for something to happen.

No credit

Code 0: Other responses.

Code 9: Missing.

This item was easy. About three-quarters of students answered correctly. The item requires students to demonstrate a broad understanding of the text by identifying its main idea. The item requires making links across the text and generalising about its overall action. The easiness of the item comes from the fact that the main idea of the text is implied and reinforced across the whole text.

QUESTION 4

According to the second last paragraph ("Before long …"), in what way are the night and a sea similar?

A. Both hide what is in them.

B. Both are noisy.

C. Both have been tamed by humans.

D. Both are dangerous to humans.

E. Both are silent.

The framework characteristics are described below:

- Situation: Personal
- Medium: Print
- Text format: Continuous
- Text type: Narration
- Aspect: Integrate and interpret: Develop an interpretation
- Question intent : Understand the point of comparison in a metaphor
- Item format: Multiple choice

Full credit

Code 1: A. Both hide what is in them.

No credit

Code 0: Other responses.

Code 9: Missing.

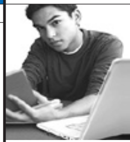
The item requires students to interpret a metaphor, although the word "metaphor" is deliberately avoided in the stem: such metalinguistic terms are likely to vary in familiarity for students from different educational backgrounds, and such metalinguistic knowledge is not part of the PISA description of reading proficiency. On the other hand, the ability to construe figurative language is considered an important constituent of interpreting texts, and particularly literary texts. It is recognised that a particular challenge for an international assessment of reading is to reflect this ability across languages and cultures. In this item, the figurative language in question uses terms ("sea" and "night") that can be regarded as universally familiar, and that have a similar connotation across cultures in the context provided by the narrative passage. The field trial results indicate that the item had robust psychometric qualities and performed similarly across countries and languages. This item demonstrates, then, that it is sometimes possible to successfully construct an item that focuses on a text's literary qualities, such as figurative language, for an international assessment. This question also demonstrates that while it is most common for multiple-choice items in PISA to have four possible response options, sometimes more than four options are given. The item was moderately difficult, with less than two-thirds of students answering it correctly.

Notes

1. The discussion in this section refers to reading in both print and digital media, unless otherwise stated.

2. This does not preclude the use of several texts in a single task, but each of the texts should be coherent in itself.

3. The hypertext link is a technique that appeared in the 1980s as a way of connecting units of information in large digital documents (Conklin, 1987; Koved and Shneiderman, 1986; Lachman, 1989; Weyer, 1982). The hypertext link or hyperlink is a piece of information (a word or phrase, or a picture or icon) that is logically connected to another piece of information (usually a page). The use of hyperlinks allows for the creation of multi-page documents with a networked structure.



References

Binkley, M. and **P. Linnakylâ** (1997), "Teaching Reading in the United States and Finland", in M. Binkley, K. Rust and T. Williams (eds.), *Reading Literacy in an International Perspective,* US Department of Education, Washington D.C.

Bruner, J. (1990), *Acts of Meaning,* Harvard University Press, Cambridge, Massachusetts.

Conklin, J. (1987), "Hypertext: An Introduction and Survey", *Computer,* Vol. 20, pp.17- 41.

Coulombe, S., J-F. Tremblay and **S. Marchand** (2004), *Literacy Scores, Human Capital, and Growth Across Fourteen OECD Countries*, Statistics Canada, Ottawa.

Council of Europe (1996), *Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference*, CC LANG (95) 5 Rev. IV, Council of Europe, Strasbourg.

Cunningham, A.E. and **K.E. Stanovich** (1998), "Early Reading Acquisition and its Relation to Reading Experience and Ability 10 Years Later", *Developmental Psychology,* Vol. 33, pp. 934-945.

Dole, J.G. Duffy, L. Roehler and D. Pearson (1991), "Moving from the Old to the New: Research on Reading Comprehension Instruction", *Review of Educational Research,* Vol. 16 (2), pp. 239-264.

Fastrez, P. (2001), "Characteristic(s) of Hypermedia and how they Relate to Knowledge", *Education Media International,* Vol. 38, pp. 101-110.

Halpern, D.F. (1989), *Thought and Knowledge: An Introduction to Critical Thinking,* Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Holloway, J.H. (1999), "Improving the Reading Skills of Adolescents", *Educational Leadership,* Vol. 57(2), pp. 80-82.

Hubbard, R. (1989), "Notes from the Underground: Unofficial Literacy in one Sixth Grade", *Anthropology and Education Quarterly,* Vol. 20, pp. 291-307.

Kirsch, I. (2001), *The International Adult Literacy Survey: Understanding What Was Measured,* Educational Testing Service, Princeton, New Jersey.

Kirsch, I. and **P.B. Mosenthal** (1990), "Exploring Document Literacy: Variables Underlying the Performance of Young Adults", *Reading Research Quarterly,* Vol. 25(1), pp. 5-30.

Koved, L. and **B. Shneiderman** (1986), "Embedded Menus: Selecting Items in Context", *Communications of the ACM,* Vol. 29(4), pp. 312-318.

Lachman, R. (1989), "Comprehension Aids for Online Reading of Expository Text", *Human Factors,* Vol. 31, pp. 1-15.

Legros, D. and **J. Crinon** (eds.) (2002), *Psychologie des apprentissages et multimedia,* Armand Colin, Paris.

Leu, D. (2007), *Expanding the Reading Literacy Framework of PISA 2009 to include Online Reading Comprehension*, unpublished manuscript.

OECD (2009), *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science,* PISA, OECD Publishing.

Pew Internet and American Life Project (2005), *Internet: The Mainstreaming of Online Life, Trends 2005,* Washington, D.C.

Reinking, D. (1994), "Electronic Literacy", *Perspectives in Reading Research,* Vol. 4.

Shetzer, H. and **M. Warschauer** (2000), "An Electronic Literacy Approach to Network-based Language Teaching", in M. Warschauer and R. Kem (eds.), *Network-based Language Teaching: Concepts and Practice*, Cambridge University Press, New York, pp. 171-185.

Smith, M.C., L. Mikulecky, M.W. Kibby and M.J. Dreher (2000), "What will be the Demands of Literacy in the Workplace in the Next Millennium?", *Reading Research Quarterly,* Vol. 35(3), pp. 378-383.

Sticht, T.G. (ed.) (1975), *Reading for Working: A Functional Literacy Anthology*, Human Resources Research Organization, Alexandria, Victoria.

Stiggins, R.J. (1982), "An Analysis of the Dimensions of Job-related Reading", *Reading World,* Vol. 82, pp. 237-247.

Sweets, R. and **A. Meates** (2004), *ICT and Low Achievers: What does PISA tell us?,* Hungarian Ministry of Education and OECD, Budapest and Paris.

Warschauer, M. (1999), *Electronic Literacies: Language Culture and Power in Online Education,* Lawrence Erlbaum Associates, Mahwah, New Jersey.

Werlich, E. (1976), *A Text Grammar of English*, Quelle and Meyer, Heidelberg.

Weyer, S.A. (1982), "The Design of a Dynamic Book for Information Search", *International Journal of Man-Machine Studies,* Vol. 17, pp. 87-107.

3

PISA 2012 Science Framework

This chapter presents the theory underlying the PISA 2012 science assessment. It begins with a definition of scientific literacy, outlines the organisation of science in PISA and sets the context for the test questions. The chapter describes the knowledge and skills at the heart of the assessment: *identifying scientific issues, explaining phenomena scientifically* and *using scientific evidence*. It then describes how knowledge and attitudes are also encompassed in the PISA definition of scientific literacy. Test questions are given as examples throughout this chapter to illustrate the classification, format and structure of the PISA science assessment.

INTRODUCTION

This framework describes and illustrates the definition of *scientific literacy* as used in PISA and sets the context for the items. Science is a minor domain in PISA 2012. The definition of the domain is unchanged since PISA 2006 when for the first time it was the major domain assessed (OECD, 2006; Bybee and McCrae, 2009), though there are some changes in terminology, which have been brought about by an attempt to better align the language used in PISA with the language used in the DeSeCo initiative (OECD, 2003).

In this framework, the term "science literacy" denotes an overarching competency comprising a set of three specific scientific competencies. A competency is more than just knowledge and skills (OECD, 2003). It includes the capacity to mobilise cognitive and non-cognitive resources in any given context. When discussing the cognitive dimensions of the specific scientific competencies, as is pertinent to the PISA science assessment in the current cycle, reference is made to the relevant scientific knowledge and skills demonstrated by students. However, the sub-scales of the PISA science scale as established in PISA 2006 (OECD, 2006) are still referred to as scientific competencies.

In keeping with its status as a minor domain in this cycle, the student questionnaire will not include items asking about students' general attitudes towards science; nor will the main assessment instrument include questions on attitudes alongside the testing of cognitive abilities and knowledge, as was the case in PISA 2006. In this revised version of the science framework for PISA 2012, like for PISA 2009, the section describing the PISA science assessment has been revised to reflect these changes, the discussion on reporting scales has been updated, and released examples from PISA 2006 have been included to illustrate the framework.

An understanding of science and technology is central to a young person's preparedness for life in modern society. It enables an individual to participate fully in a society in which science and technology play a significant role. This understanding also empowers individuals to participate appropriately in the determination of public policy where issues of science and technology impact on their lives. An understanding of science and technology contributes significantly to the personal, social, professional and cultural lives of everyone.

A large proportion of the situations, problems and issues encountered by individuals in their daily lives require some understanding of science and technology before they can be fully understood or addressed. Science and technology related issues confront individuals at personal, community, national and even global levels. Therefore, national leaders should be encouraged to ask about the degree to which all individuals in their respective countries are prepared to deal with these issues. A critical aspect of this is how young people respond to scientific questions when they emerge from school. An assessment at age 15 provides an early indication of how students may respond later in life to the diverse array of situations that involve science and technology.

As the basis for an international assessment of 15-year-old students, it seems reasonable, therefore, to ask: "What is it important for citizens to know, value, and be able to do in situations involving science and technology?" Answering this question establishes the basis for an assessment of students with regards to how their knowledge, values and abilities today relate to what they will need in the future. Central to the answer are the competencies that lie at the heart of the PISA science assessment. These ask how well students:

- *identify scientific issues;*
- *explain phenomena scientifically*; and
- *use scientific evidence*.

These competencies require students to demonstrate, on the one hand, knowledge and cognitive abilities, and on the other, attitudes, values and motivations, as they meet and respond to science-related issues.

The issue of identifying what citizens should know, value and be able to do in situations involving science and technology, seems simple and direct. However, doing so raises questions about scientific understanding and does not imply mastery of all scientific knowledge. This framework is guided by reference to what citizens require. As citizens, what knowledge is most appropriate? An answer to this question certainly includes basic concepts of the science disciplines, but that knowledge must be used in contexts that individuals encounter in life. In addition, people often encounter situations that require some understanding of science as a process that produces knowledge and proposes explanations about the natural world.[1] Further, they should be aware of the complementary relationships between science and technology, and how science-based technologies pervade and influence the nature of modern life.

What is important for citizens to value about science and technology? An answer should include the role and contributions to society of science and of science-based technology, and their importance in many *personal*, *social*, and *global* contexts. Accordingly, it seems reasonable to expect individuals to have an interest in science, to support the process of scientific enquiry and to act responsibly towards natural resources and the environment.

What is important for individuals to be able to do that is science-related? People often have to draw appropriate conclusions from evidence and information given to them; they have to evaluate claims made by others on the basis of the evidence put forward and they have to distinguish personal opinion from evidence-based statements. Often the evidence involved is scientific, but science has a more general role to play as well since it is concerned with rationality in testing ideas and theories against evidence. Of course this does not deny that science includes creativity and imagination, attributes that have always played a central part in advancing human understanding of the world.

Can citizens distinguish claims that are scientifically sound from those that are not? Ordinary citizens are generally not called on to judge the worth of major theories or potential advances in science. But they do make decisions based on the facts in advertisements, evidence in legal matters, information about their health, and issues concerning local environments and natural resources. An educated person should be able to distinguish the kinds of questions that can be answered by scientists and the kinds of problems that can be solved by science-based technologies from those that cannot be answered in these ways.

DEFINING SCIENTIFIC LITERACY

Current thinking about the desired outcomes of science education emphasises scientific knowledge (including knowledge of the scientific approach to enquiry) and an appreciation of science's contribution to society. These outcomes require an understanding of important concepts and explanations of science, and of the strengths and limitations of science in the world. They imply a critical stance and a reflective approach to science (Millar and Osborne, 1998).

Such goals provide an orientation and emphasis for the science education of all people (Fensham, 1985). The competencies assessed in PISA are broad and include aspects that relate to personal utility, social responsibility, and the intrinsic and extrinsic value of scientific knowledge.

The above discussion frames a central point of the PISA science assessment: the assessment should focus on scientific competencies that clarify what 15-year-old students know, value and are able to do within reasonable and appropriate *personal*, *social*, and *global* contexts. This perspective differs from one grounded exclusively in school science programmes and extensively based only on the disciplines of science; but it includes problems situated in educational contexts and also in professional ones, and recognises the essential place of the knowledge, methods, attitudes, and values that define scientific disciplines (Bybee, 1997a; Fensham, 2000; Gräber and Bolte, 1997; Mayer, 2002; Roberts, 1983; UNESCO, 1993).

PISA is concerned with both the cognitive and affective aspects of students' competencies in science. The cognitive aspects include students' knowledge and their capacity to use this knowledge effectively, as they carry out certain cognitive processes that are characteristic of science and scientific enquiries of *personal*, *social*, and *global* relevance. In assessing scientific competencies, PISA is concerned with issues to which scientific knowledge can contribute and which will involve students, either now or in the future, in making decisions. From the point of view of their scientific competencies, students respond to such issues in terms of their understanding of relevant scientific knowledge, their ability to access and evaluate information, their ability to interpret evidence bearing on the issue and their ability to identify the scientific and technological aspects of the issue (Koballa et al., 1997; Law, 2002). PISA also is concerned with non-cognitive aspects: how students respond affectively. Attitudinal aspects of their response engage their interest, sustain their support, and motivate them to take action (Schibeci, 1984).

Box 3.1 **Scientific knowledge: PISA terminology**

The term "scientific knowledge" is used throughout this framework to refer to both *knowledge of science* and *knowledge about science*. *Knowledge of science* refers to knowledge of the natural world across the major fields of physics, chemistry, biological science, Earth and space science, and science-based technology. *Knowledge about science* refers to knowledge of the means ("scientific enquiry") and goals ("scientific explanations") of science.

The PISA science assessment encompasses a continuum of scientific knowledge and the cognitive abilities associated with scientific enquiry, incorporates multiple dimensions, and addresses the relationships between science and technology. It provides an assessment of students' scientific literacy by assessing their capacity to use scientific knowledge (Bybee, 1997b; Fensham, 2000; Law, 2002; Mayer and Kumano, 2002).

Box 3.2 **PISA scientific literacy**

For the purposes of PISA, scientific literacy refers to an individual's:

- Scientific knowledge and use of that knowledge to identify questions, acquire new knowledge, explain scientific phenomena and 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.
- Willingness to engage in science-related issues, and with the ideas of science, as a reflective citizen.

Explanation of the definition

The following remarks clarify the definition of scientific literacy as defined for the purposes of PISA.

Using the term "scientific literacy" rather than "science" underscores the importance that the PISA science assessment places on the application of scientific knowledge in the context of life situations, compared with the simple reproduction of traditional school science knowledge. The functional use of knowledge requires the application of those processes that are characteristic of science and scientific enquiry (here termed the scientific competencies) and is regulated by the individual's appreciation, interest, values, and action relative to scientific matters. A student's ability to carry out the scientific competencies involves both *knowledge of science* and an understanding of the characteristics of science as a way of acquiring knowledge (i.e. *knowledge about science*). The definition also recognises that the disposition to carry out these competencies depends upon an individual's attitudes towards science and a willingness to engage in science-related issues.

Knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena and to draw evidence-based conclusions

"Knowledge" for this definition implies far more than the ability to recall information, facts, and names. The definition includes *knowledge of science* (knowledge about the natural world) and *knowledge about science* itself. *Knowledge of science* includes understanding fundamental scientific concepts and theories; *knowledge about science* includes understanding the nature of science as a human activity and the power and limitations of scientific knowledge. The questions to be identified are those that can be answered by scientific enquiry, again requiring *knowledge about science* as well as scientific knowledge of the specific topics involved. Of significance is that individuals must often acquire new knowledge not through their own scientific investigations, but through resources such as libraries and the Internet. Drawing evidence-based conclusions means knowing, selecting and evaluating information and data, while recognising that there is often not sufficient information to draw definite conclusions, thus making it necessary to speculate cautiously and consciously about the information that is available.

Characteristic features of science as a form of human knowledge and enquiry

As expressed here, being scientifically literate implies that students should have some understanding of how scientists obtain data and propose explanations, and recognise key features of scientific investigations and the types of answers one can reasonably expect from science. For example, scientists use observations and experiments to gather data about objects, organisms and events in the natural world. The data are used to propose explanations that become public knowledge and may be used in various forms of human activity. Some key features of science include: the collection and use of data – data collection is guided by ideas and concepts (sometimes stated as hypotheses) and includes issues of relevance, context and accuracy; the tentative nature of knowledge claims; an openness to sceptical review; the use of logical arguments; and, the obligation to make connections to current and historical knowledge, and to report the methods and procedures used in obtaining evidence.

How science and technology shape our material, intellectual, and cultural environments

The key points in this statement include the idea that science is a human endeavour, one that influences our societies and us as individuals. Further, technological development is also a human endeavour (Fleming, 1989). Although science and technology differ in aspects of their purposes, processes, and products, they are also closely related and, in many respects, complementary. In this regard, the definition of scientific literacy as used here includes the nature of science and of technology and their complementary relationships. As individuals we make decisions through public policies that influence the directions of science and technology. Science and technology play paradoxical roles in society as they propose answers to questions and provide solutions to problems, but may also create new questions and problems.

Willingness to engage in science-related issues and with the ideas of science as a reflective citizen

The meaning conveyed in the first part of this statement, 'willingness to engage in science-related issues', is wider than taking note and taking action as required; it implies having continuing interest in, having opinions about and participating in current and future science-based issues. The second part of the statement, 'with the ideas of science as a reflective citizen', covers various aspects of attitudes and values that individuals may have towards science. The whole phrase implies a person who has an interest in scientific topics, thinks about science-related issues, is concerned about issues of technology, resources and the environment, and reflects on the importance of science in personal and social perspectives.

Inevitably, scientific competencies draw upon reading and mathematical competencies (Norris and Phillips, 2003). For example, aspects of mathematical competencies are required in data interpretation contexts. Similarly, reading literacy is necessary when a student is demonstrating an understanding of scientific terminology. The intersection of these other domains with the PISA definition and assessment of science cannot be avoided; however, at the core of each assessment task there should be aspects that relate unambiguously to science competency.

ORGANISING THE DOMAIN

The definition of the science domain proposed here provides for a continuum in which individuals are deemed to be more or less scientifically literate; they are not regarded as either scientifically literate or scientifically illiterate (Bybee, 1997a; 1997b). So, for example, the student with less developed scientific literacy might be able to recall simple scientific factual knowledge and to use common scientific knowledge in drawing or evaluating conclusions. A student with more developed scientific literacy will demonstrate the ability to create and use conceptual models to make predictions and give explanations, analyse scientific investigations, relate data as evidence, evaluate alternative explanations of the same phenomena, and communicate conclusions with precision.

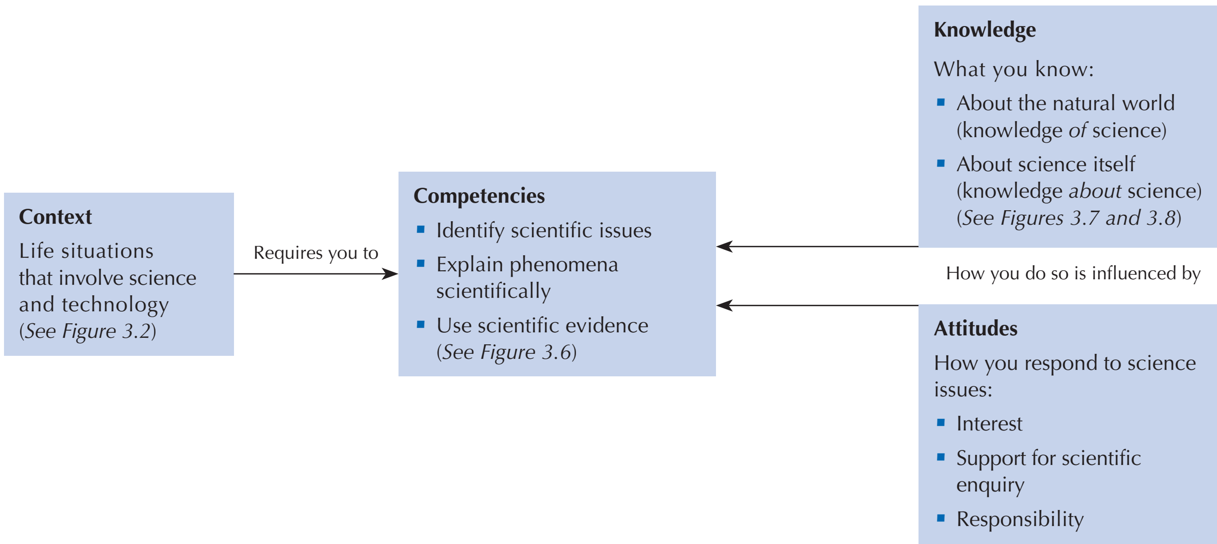
For assessment purposes, the PISA definition of *scientific literacy* may be characterised as consisting of four interrelated aspects (see Figure 3.1):

- Context: recognising life situations involving science and technology.
- Knowledge: understanding the natural world on the basis of scientific knowledge that includes both knowledge of the natural world, and *knowledge about science* itself.
- Competencies: demonstrating scientific competencies that include *identifying scientific issues, explaining phenomena scientifically*, and *using scientific evidence*.
- Attitudes: indicating an interest in science, support for scientific enquiry, and motivation to act responsibly towards, for example, natural resources and environments.

The following sections restate and elaborate these interrelated aspects. In highlighting these aspects, the PISA science framework has ensured that the focus of the assessment is upon the outcomes of science education. Several questions have guided the establishment of the PISA science framework. They are:

- What contexts would be appropriate for assessing 15-year-old students?
- What competencies might we reasonably expect 15-year-old students to demonstrate?
- What knowledge might we reasonably expect 15-year-old students to demonstrate?
- What attitudes might we reasonably expect 15-year-old students to demonstrate?

■ Figure 3.1 ■
Framework for the PISA science assessment



Situations and context

An important aspect of the PISA science assessment is engagement with science in a variety of situations. In dealing with scientific issues, the choice of methods and representations is often dependent on the situations in which the issues are presented.

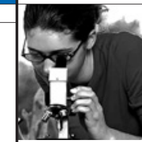
The situation is the part of the student's world in which the tasks are placed. Assessment items are framed in situations of general life and not limited to life in school. In the PISA science assessment, the focus of the items is on situations relating to the self, family and peer groups (*personal*), to the community (*social*) and to life across the world (*global*). A further type of setting, appropriate to some topics, is the historical one, in which understanding of the advances in scientific knowledge can be assessed.

PISA assesses important scientific knowledge relevant to the science education curricula of participating countries without being constrained to the common aspects of participants' national curricula. The assessment does this by requiring evidence of the successful use of science knowledge and skills in important situations reflecting the world and in accordance with the PISA focus. This, in turn, involves the application of selected knowledge about the natural world, and about science itself, and evaluation of students' attitudes towards scientific matters.

Figure 3.2 lists the main applications of science that are involved within *personal*, *social*, and *global* settings as the contexts, or specific situations, for assessment exercises. However, other settings (e.g. technological, historical) and areas of application are also used. The areas of application are: "health", "natural resources", "the environment", "hazards", and "the frontiers of science and technology". They are the areas in which science has particular value for individuals and communities in enhancing and sustaining quality of life, and in the development of public policy.

The PISA science assessment is not an assessment of contexts. It assesses competencies, knowledge and attitudes as these are presented or relate to contexts. In selecting the contexts, it is important to keep in mind that the purpose of the assessment is to assess scientific competencies, understandings, and attitudes that students have acquired by the end of the compulsory years of schooling.

PISA items are arranged in groups (units) based around a common stimulus that establishes the context for the items. The contexts used are chosen in the light of relevance to students' interests and lives. The items are developed keeping in mind linguistic and cultural differences in participating countries.



■ Figure 3.2 ■

Contexts for the PISA science assessment

	Personal (self, family and peer groups)	**Social** (the community)	**Global** (life across the world)
Health	Maintenance of health, accidents, nutrition	Control of disease, social transmission, food choices, community health	Epidemics, spread of infectious diseases
Natural resources	Personal consumption of materials and energy	Maintenance of human populations, quality of life, security, production and distribution of food, energy supply	Renewable and non-renewable energy sources, natural systems, population growth, sustainable use of species
Environment	Environmentally friendly behaviour, use and disposal of materials	Population distribution, disposal of waste, environmental impact, local weather	Biodiversity, ecological sustainability, control of pollution, production and loss of soil
Hazard	Natural and human-induced, decisions about housing	Rapid changes (earthquakes, severe weather), slow and progressive changes (coastal erosion, sedimentation), risk assessment	Climate change, impact of modern warfare
Frontiers of science and technology	Interest in science's explanations of natural phenomena, science-based hobbies, sport and leisure, music and personal technology	New materials, devices and processes, genetic modification, weapons technology, transport	Extinction of species, exploration of space, origin and structure of the universe

Illustrative PISA science items

In this section, three examples of science units from the PISA 2006 assessment are presented. They are referred to throughout the remainder of the chapter to illustrate the variety of contexts involved, the scientific competencies and areas of scientific knowledge addressed by PISA science items, and the item types (formats) employed. In addition, the scoring guide for each item is shown (for a description of proficiency levels, see Figure 3.10).

ACID RAIN

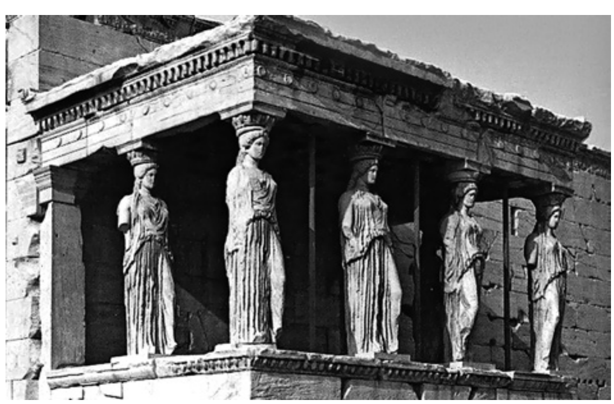
In this example, the stimulus material is a photograph of statues on the Acropolis in Athens, together with a brief statement explaining that the original statues were moved inside the museum of the Acropolis due to their deterioration from acid rain. The area of application is "Hazards" within *personal* and *social* settings.

■ Figure 3.3 ■

Items for the unit *ACID RAIN*

Below is a photo of statues called Caryatids that were built on the Acropolis in Athens more than 2500 years ago. The statues are made of a type of rock called marble. Marble is composed of calcium carbonate.

In 1980, the original statues were transferred inside the museum of the Acropolis and were replaced by replicas. The original statues were being eaten away by acid rain.



QUESTION 1

Normal rain is slightly acidic because it has absorbed some carbon dioxide from the air. Acid rain is more acidic than normal rain because it has absorbed gases like sulfur oxides and nitrogen oxides as well.

Where do these sulfur oxides and nitrogen oxides in the air come from?

Full credit (Level 3: 506)

Any one of car exhausts, factory emissions, *burning* fossil fuels such as oil and coal, gases from volcanoes or other similar things *OR* Responses that include an incorrect as well as a correct source of the pollution *OR* Responses that refer to "pollution" but do not give a source of pollution that is a significant cause of acid rain.

The effect of acid rain on marble can be modelled by placing chips of marble in vinegar overnight. Vinegar and acid rain have about the same acidity level. When a marble chip is placed in vinegar, bubbles of gas form. The mass of the dry marble chip can be found before and after the experiment.

QUESTION 2

A marble chip has a mass of 2.0 grams before being immersed in vinegar overnight. The chip is removed and dried the next day.

What will the mass of the dried marble chip be?

A. Less than 2.0 grams

B. Exactly 2.0 grams

C. Between 2.0 and 2.4 grams

D. More than 2.4 grams

Full credit (Level 2: 460)

A. Less than 2.0 grams

QUESTION 3

Students who did this experiment also placed marble chips in pure (distilled) water overnight.

Explain why the students included this step in their experiment.

Full credit (Level 6: 717)

To show that the acid (vinegar) is necessary for the reaction.

Partial credit (Level 3: 513)

To compare with the test of vinegar and marble, but it is not made clear that this is being done to show that the acid (vinegar) is necessary for the reaction.

GREENHOUSE

This unit deals with the increase of the average temperature of the Earth's atmosphere. The stimulus material consists of a short text introducing the term "Greenhouse effect" and includes graphical information on the average temperature of the Earth's atmosphere and the carbon dioxide emission on the Earth over time.

The area of application is "Environment" within a *global* setting.

■ Figure 3.4 ■

Items for the unit *GREENHOUSE*

Read the texts and answer the questions that follow.

THE GREENHOUSE EFFECT: FACT OR FICTION?

Living things need energy to survive. The energy that sustains life on the Earth comes from the Sun, which radiates energy into space because it is so hot. A tiny proportion of this energy reaches the Earth.

The Earth's atmosphere acts like a protective blanket over the surface of our planet, preventing the variations in temperature that would exist in an airless world.

Most of the radiated energy coming from the Sun passes through the Earth's atmosphere. The Earth absorbs some of this energy, and some is reflected back from the Earth's surface. Part of this reflected energy is absorbed by the atmosphere.

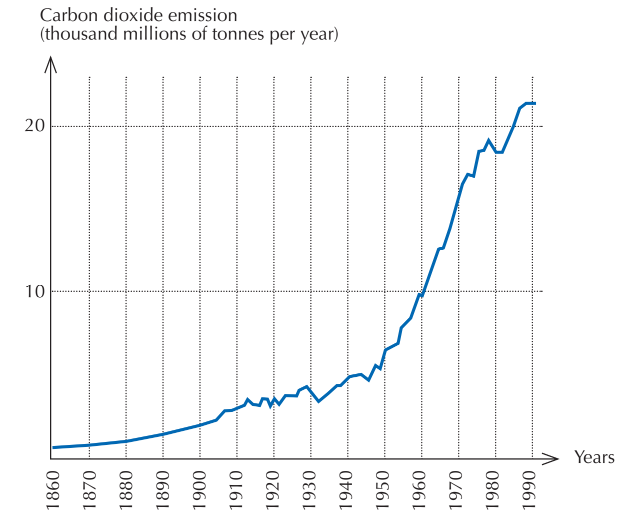
As a result of this the average temperature above the Earth's surface is higher than it would be if there were no atmosphere. The Earth's atmosphere has the same effect as a greenhouse, hence the term greenhouse effect.

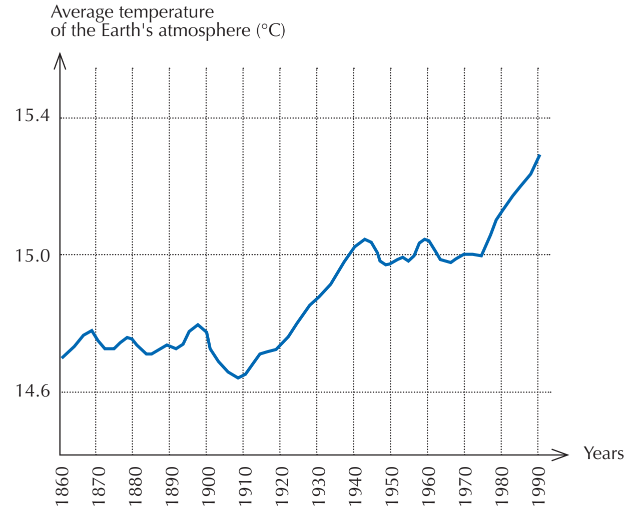
The greenhouse effect is said to have become more pronounced during the twentieth century.

It is a fact that the average temperature of the Earth's atmosphere has increased. In newspapers and periodical the increased carbon dioxide emission is often stated as the main source of the temperature rise in the twentieth century.

A student named André becomes interested in the possible relationship between the average temperature of the Earth's atmosphere and the carbon dioxide emission on the Earth.

In a library he comes across the following two graphs.





André concludes from these two graphs that it is certain that the increase in the average temperature of the Earth's atmosphere is due to the increase in the carbon dioxide emission.

QUESTION 1

What is it about the graphs that supports André's conclusion?

Full credit (Level 3: 529)

Refers to the increase of both (average) temperature and carbon dioxide emission OR Refers (in general terms) to a positive relationship between temperature and carbon dioxide emission.

QUESTION 2

Another student, Jeanne, disagrees with André's conclusion. She compares the two graphs and says that some parts of the graphs do not support his conclusion.

Give an example of a part of the graphs that does not support André's conclusion. Explain your answer.

Full credit (Level 5: 659)

Refers to one particular part of the graphs in which the curves are not both descending or both climbing and gives the corresponding explanation.

Partial credit (Level 4: 568)

Mentions a correct period, without any explanation OR Mentions only one particular year (not a period of time), with an acceptable explanation OR Gives an example that doesn't support André's conclusion but makes a mistake in mentioning the period OR Refers to differences between the two curves, without mentioning a specific period OR Refers to an irregularity in one of the graphs OR Indicates difference in the graphs, but explanation is poor.

QUESTION 3

André persists in his conclusion that the average temperature rise of the Earth's atmosphere is caused by the increase in the carbon dioxide emission. But Jeanne thinks that his conclusion is premature. She says: "Before accepting this conclusion you must be sure that other factors that could influence the greenhouse effect are constant".

Name one of the factors that Jeanne means.

Full credit (Level 6: 709)

Gives a factor referring to the energy/radiation coming from the Sun OR Gives a factor referring to a natural component or a potential pollutant.

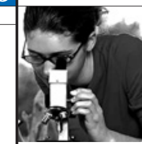
PHYSICAL EXERCISE

This unit is concerned with the effect of physical exercise on personal health.

■ Figure 3.5 ■

Items for the unit *PHYSICAL EXERCISE*

Regular but moderate physical exercise is good for our health.



QUESTION 1

What are the advantages of regular physical exercise? Circle "Yes" or "No" for each statement.

Is this an advantage of regular physical exercise?	Yes or No ?
Physical exercise helps prevent heart and circulation illnesses.	Yes / No
Physical exercise leads to a healthy diet.	Yes / No
Physical exercise helps to avoid becoming overweight.	Yes / No

Full credit (Level 3: 545)

All three correct: Yes, No, Yes in that order.

QUESTION 2

What happens when muscles are exercised? Circle "Yes" or "No" for each statement.

Does this happen when muscles are exercised?	Yes or No ?
Muscles get an increased flow of blood.	Yes / No
Fats are formed in the muscles.	Yes / No

Full credit (Level 1: 386)

Both correct: Yes, No in that order.

QUESTION 3

Why do you have to breathe more heavily when you're doing physical exercise than when your body is resting?

....

Full credit (Level 4: 583)

To remove increased levels of carbon dioxide and to supply more oxygen to your body OR To remove increased levels of carbon dioxide from your body OR To supply more oxygen to your body, but not both.

Scientific competencies

The PISA science assessment gives priority to the competencies listed in Figure 3.6: the ability to identify scientifically-oriented issues; describe, explain or predict phenomena based on scientific knowledge; interpret evidence and conclusions, and use scientific evidence to make and communicate decisions. Demonstrating these competencies in the PISA assessment involves applying scientific knowledge – both *knowledge of science* and *knowledge about science* itself as a form of knowledge and an approach to enquiry.

■ Figure 3.6 ■

PISA scientific competencies

Identifying scientific issues
- Recognising issues that are possible to investigate scientifically
- Identifying keywords to search for scientific information
- Recognising the key features of a scientific investigation

Explaining phenomena scientifically
- Applying knowledge of science in a given situation
- Describing or interpreting phenomena scientifically and predicting changes
- Identifying appropriate descriptions, explanations, and predictions

Using scientific evidence
- Interpreting scientific evidence and making and communicating conclusions
- Identifying the assumptions, evidence and reasoning behind conclusions
- Reflecting on the societal implications of science and technological developments

Some cognitive processes have special meaning and relevance for *scientific literacy*. Among the cognitive processes that are implied in the scientific competencies are: inductive reasoning (reasoning from detailed facts to general principles) and deductive reasoning (reasoning from the general to the particular), critical and integrated thinking, transforming representations (e.g. data to tables, tables to graphs), constructing and communicating arguments and explanations based on data, thinking in terms of models, and using mathematical processes, knowledge and skills.

Justification for an emphasis on the scientific competencies of Figure 3.6 in PISA rests on the importance of these competencies for scientific investigation. They are grounded in logic, reasoning, and critical analysis. An elaboration of the scientific competencies follows, including references to how they are assessed in the science examples presented in the previous section.

Identifying scientific issues

It is important to be able to distinguish scientific issues and content from other forms of issues. Importantly, scientific issues must lend themselves to answers based on scientific evidence. The competency *identifying scientific issues* includes recognising questions that it would be possible to investigate scientifically in a given situation and identifying keywords to search for scientific information on a given topic. It also includes recognising key features of a scientific investigation: for example, what things should be compared, what variables should be changed or controlled, what additional information is needed, or what action should be taken so that relevant data can be collected.

Identifying scientific issues requires students to possess *knowledge about science* itself, and may also draw, to varying degrees, on their *knowledge of science*. For example, Question 3 of *ACID RAIN* requires students to answer a question about the control in a scientific investigation. Students must compare an acid (vinegar) reaction to possible reactions with pure water to be sure that acid is the cause of the reaction.

Explaining phenomena scientifically

Demonstrating the competency *explaining phenomena scientifically* involves applying appropriate *knowledge of science* in a given situation. The competency includes describing or interpreting phenomena and predicting changes, and may involve recognising or identifying appropriate descriptions, explanations, and predictions. An example of a PISA item that requires students to explain phenomena scientifically is Question 1 of *ACID RAIN*, where students must explain the origin of sulfur oxides and nitrogen oxides in the air. Other examples are: Question 3 of *GREENHOUSE*, which requires students to identify factors causing the average temperature rise of the Earth, and Question 3 of *PHYSICAL EXERCISE* which asks students to apply their knowledge of the human respiratory system.

Using scientific evidence

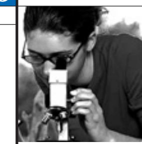
The competency *using scientific evidence* includes accessing scientific information and producing arguments and conclusions based on scientific evidence (Kuhn, 1992; Osborne et al., 2001). The required response can involve *knowledge about science* or *knowledge of science* or both. Question 2 of *ACID RAIN* requires students to use the information provided to form a conclusion about the effects of vinegar on marble, a simple model for the influence of acid rain on marble. Other examples are Questions 1 and 2 of *GREENHOUSE*, both of which require students to interpret evidence presented in two graphs.

The competency also involves: selecting from alternative conclusions in relation to evidence; giving reasons for or against a given conclusion in terms of the process by which the conclusion was derived from the data provided; and identifying the assumptions made in reaching a conclusion. Reflecting on the societal implications of scientific or technological developments is another aspect of this competency.

Students may be required to express their evidence and decisions to a specified audience, through their own words, diagrams or other representations as appropriate. In short, students should be able to present clear and logical connections between evidence and conclusions or decisions.

Scientific knowledge

As previously noted (see Box 3.1), scientific knowledge refers to both *knowledge of science* (knowledge about the natural world) and *knowledge about science* itself.



Knowledge of science

Given that only a sample of students' *knowledge of science* can be assessed in any one PISA assessment, it is important that clear criteria are used to guide the selection of knowledge that is assessed. Moreover, the objective of PISA is to describe the extent to which students can apply their knowledge in contexts of relevance to their lives. Accordingly, the assessed knowledge will be selected from the major fields of physics, chemistry, biology, Earth and space science, and technology according to the following criteria:

- The relevance to real-life situations: scientific knowledge differs in the degree to which it is useful in the life of individuals.
- The knowledge selected represents important scientific concepts and thus has enduring utility.
- The knowledge selected is appropriate to the developmental level of 15-year-old students.

Figure 3.7 shows the *knowledge of science* categories and examples of content selected by applying these criteria. This knowledge is required for understanding the natural world and for making sense of experiences in *personal*, *social* and *global* situations. The framework uses the term "systems" instead of "sciences" in the descriptors of the major fields to convey the idea that citizens have to apply their understanding of concepts from the physical and life sciences, Earth and space science, and technology, in situations that interact in more or less united ways.

■ Figure 3.7 ■

PISA categories of knowledge of science

Physical systems
- Structure of matter (e.g. particle model, bonds)
- Properties of matter (e.g. changes of state, thermal and electrical conductivity)
- Chemical changes of matter (e.g. reactions, energy transfer, acids/bases)
- Motions and forces (e.g. velocity, friction)
- Energy and its transformation (e.g. conservation, dissipation, chemical reactions)
- Interactions of energy and matter (e.g. light and radio waves, sound and seismic waves)

Living systems
- Cells (e.g. structures and function, DNA, plant and animal)
- Humans (e.g. health, nutrition, subsystems [i.e. digestion, respiration, circulation, excretion, and their relationship], disease, reproduction)
- Populations (e.g. species, evolution, biodiversity, genetic variation)
- Ecosystems (e.g. food chains, matter and energy flow)
- Biosphere (e.g. ecosystem services, sustainability)

Earth and space systems
- Structures of Earth systems (e.g. lithosphere, atmosphere, hydrosphere)
- Energy in Earth systems (e.g. sources, global climate)
- Change in Earth systems (e.g. plate tectonics, geochemical cycles, constructive and destructive forces)
- Earth's history (e.g. fossils, origin and evolution)
- Earth in space (e.g. gravity, solar systems)

Technology systems
- Role of science-based technology (e.g. solve problems, help humans meet needs and wants, design and conduct investigations)
- Relationships between science and technology (e.g. technologies contribute to scientific advancement)
- Concepts (e.g. optimisation, trade-offs, cost, risk, benefit)
- Important principles (e.g. criteria, constraints, innovation, invention, problem solving)

The examples listed in Figure 3.7 convey the meanings of the categories; there is no attempt to list comprehensively all the knowledge that could be related to each of the *knowledge of science* categories.

Question 2 of *ACID RAIN* assesses students' *knowledge of science* in the category "Physical systems".

Question 3 of *GREENHOUSE* is concerned with students' knowledge of "Earth and space systems"; and Questions 1, 2 and 3 of *PHYSICAL EXERCISE* assess students' knowledge of "Living systems".

Knowledge about science

Figure 3.8 displays the categories and examples of content for *knowledge about science*. The first category, "Scientific enquiry," centres on enquiry as the central process of science and the various components of that process. The second category, closely related to enquiry, is "Scientific explanations". "Scientific explanations" are the results of "Scientific enquiry". One can think of enquiry as the means of science (how scientists get data) and explanations as the goals of science (how scientists use data). The examples listed in Figure 3.8 convey the general meanings of the categories; there is no attempt to list comprehensively all the knowledge that could be related to each category.

■ Figure 3.8 ■

PISA categories of knowledge about science

Scientific enquiry

- Origin (e.g. curiosity, scientific questions)
- Purpose (e.g. to produce evidence that helps answer scientific questions, current ideas/models/theories guide enquiries)
- Experiments (e.g. different questions suggest different scientific investigations, design)
- Data type (e.g. quantitative [measurements], qualitative [observations])
- Measurement (e.g. inherent uncertainty, replicability, variation, accuracy/precision in equipment and procedures)
- Characteristics of results (e.g. empirical, tentative, testable, falsifiable, self-correcting)

Scientific explanations

- Types (e.g. hypothesis, theory, model, law)
- Formation (e.g. data representation, role of extant knowledge and new evidence, creativity and imagination, logic)
- Rules (e.g. must be logically consistent; based on evidence, historical and current knowledge)
- Outcomes (e.g. produce new knowledge, new methods, new technologies; lead to new questions and investigations)

Question 3 of *ACID RAIN* is an example of *knowledge about science* in the category "Scientific enquiry" but assumes some *knowledge of science* (category "Physical systems") that students can be expected to possess. The question requires students to identify the possible purposes for a control of an investigation (competency: *Identifying Scientific Issues*).

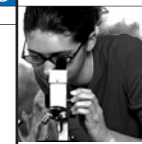
Questions 1 and 2 of *GREENHOUSE* are *knowledge about science* items. Both of these items belong to the category "Scientific explanations". In Question 1, students must interpret evidence presented in two graphs and argue that the graphs together support an explanation that an increase in the Earth's average temperature is due to an increase in carbon dioxide emissions. Question 2 asks students to use evidence from the same graphs to support a different conclusion.

Attitudes towards science

Individuals' attitudes play a significant role in their interest and response to science and technology in general and to issues that affect them in particular. One goal of science education is for students to develop attitudes that make them likely to attend to scientific issues and subsequently to acquire and apply scientific and technological knowledge for *personal*, *social*, and *global* benefit.

PISA attention to attitudes towards science is based on the belief that a person's scientific literacy includes certain attitudes, beliefs, motivational orientations, sense of self-efficacy, values, and ultimate actions. This is supported by and builds upon Klopfer's (1976) structure for the affective domain in science education, as well as reviews of attitudinal research (for example, Gardner, 1975, 1984; Gauld and Hukins, 1980; Blosser, 1984; Laforgia, 1988; Osborne et al., 2003; Schibeci, 1984) and research into students' attitudes towards the environment (for example, Bogner and Wiseman, 1999; Eagles and Demare, 1999; Weaver, 2002; Rickinson, 2001).

In PISA 2006, when science was the major domain assessed, an assessment of students' attitudes and values was included using the student questionnaire and through contextualised questions posed immediately after the test questions in many units (OECD, 2006). These contextualised questions were related to the issues addressed in the test questions. However, since science constitutes a minor part of the assessment in PISA 2012, the assessment will not contain any contextualised (embedded) attitudinal items.



ASSESSING SCIENTIFIC LITERACY

Test characteristics

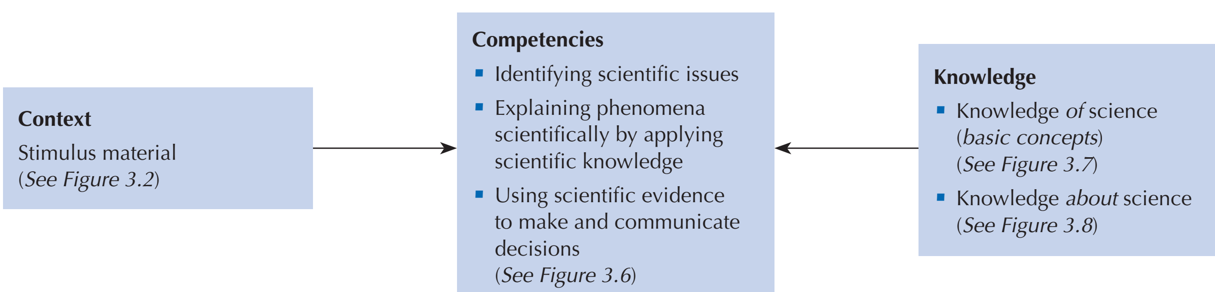
In accordance with the PISA definition of scientific literacy, test questions (items) require the use of the scientific competencies (see Figure 3.6) within a context (see Figure 3.2). This involves the application of scientific knowledge (see Figures 3.7 and 3.8).

Figure 3.9 is a variation of Figure 3.1 that presents the basic components of the PISA framework for science assessment in a way that can be used to relate the framework with the structure and the content of assessment units. Figure 3.9 may be used both synthetically as a tool to plan assessment exercises, and analytically as a tool to study the results of standard assessment exercises. As a starting point to construct assessment units, we could consider the contexts that would serve as stimulus material, the scientific competencies required to respond to the questions or issues, or the scientific knowledge central to the exercise.

A test unit is comprised of a group of independently scored questions (items) of various types, accompanied by stimulus material that establishes the context for the items. Many different types of stimulus are used, often in combination, to establish the context, including passages of text, photographs, tables, graphs, and diagrams, often in combination. The range of stimulus material is illustrated by the three units included in this chapter. *GREENHOUSE* has an extensive stimulus comprised of half a page of text and two graphs, whereas the stimulus of *PHYSICAL EXERCISE* is atypical in its brevity and reliance on visual suggestion.

■ Figure 3.9 ■

A tool for constructing and analysing assessment units and items



The reason PISA employs this unit structure is to facilitate the use of contexts that are as relevant as possible, and that reflect the complexity of real situations, while making efficient use of testing time. Using situations about which several questions can be posed, rather than asking separate questions about a larger number of different situations, reduces the overall time required for a student to become familiar with the material relating to each question. However, the need to make each scored point independent of others within a unit needs to be taken into account. It is also necessary to recognise that, because this approach reduces the number of different assessment contexts, it is important to ensure that there is an adequate range of contexts so that bias due to the choice of contexts is minimised.

PISA 2012 science test units incorporate up to four cognitive items that assess students' scientific competencies. Each item involves the predominant use of the skills involved in one of the scientific competencies, and primarily requires *knowledge of science* or *knowledge about science*. In most cases, more than one competency and more than one knowledge category are assessed (by different items) in this way within a unit.

Four types of items are used to assess the competencies and scientific knowledge identified in the framework: simple multiple-choice items, closed constructed-response items, complex multiple-choice items, and open constructed-response items. About one-third of the items are simple multiple-choice items, like Question 2 of *ACID RAIN*, which require the selection of a single response from four options. Another third of the items either require closed constructed-responses, or are complex multiple-choice items. Questions 1 and 2 of *PHYSICAL EXERCISE*, which require students to respond to a series of related "Yes/No" questions, are typical complex multiple-choice items. The remaining third of the items are open constructed-response items, like the remaining questions in *ACID RAIN* and *PHYSICAL EXERCISE* and the three items in *GREENHOUSE*. These require a relatively extended written or drawn response from students.

While the majority of items are dichotomously scored (that is, responses are either given all credit or no credit), some of the open constructed-response items allow for partial credit, and give students credit for having a partially correct answer. The categories "Full credit", "Partial credit" and "No credit" divide students' responses into three groups in terms of the extent to which the students demonstrate the ability to answer a question. A "Full credit" response requires a student to show a level of understanding of the topic that is appropriate for a scientifically literate 15-year-old student. Less sophisticated or correct responses may qualify for "Partial credit", with completely incorrect, irrelevant or missing responses being assigned "No credit". Question 3 of *ACID RAIN* and Question 2 of *GREENHOUSE* are partial credit items.

The need for students to possess a degree of reading literacy in order to understand and answer written questions on science raises an issue of the level of reading literacy required. Stimulus material and questions use language that is as clear, simple and brief as possible, while still conveying the appropriate meaning. The number of concepts introduced per paragraph is limited and questions that require too high a level of reading, or mathematics, are avoided.

Science assessment structure

Each PISA assessment must include an appropriate balance of items assessing scientific knowledge and competencies. Table 3.1 shows the distribution of score points among the *knowledge of science* and *knowledge about science* categories, expressed as percentages of the total score points, for both PISA 2006 (when science was the major domain) and PISA 2012.

Table 3.1
Approximate distribution of score points in science, by knowledge

	Percentage of score points	
Knowledge *of* science	**PISA 2006**	**PISA 2012**
Physical systems	17	13
Living systems	20	16
Earth and space systems	10	12
Technological systems	8	9
Subtotal	55	50
Knowledge *about* science		
Scientific enquiry	23	23
Scientific explanations	22	27
Subtotal	45	50
Total	**100**	**100**

The corresponding distributions for the scientific competencies are given in Table 3.2

Table 3.2
Approximate distribution of score points in science, by scientific competencies

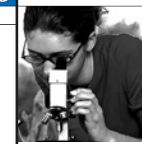
	Percentage of score points	
Scientific competencies	**PISA 2006**	**PISA 2012**
Identifying scientific issues	22	23
Explaining phenomena scientifically	46	41
Using scientific evidence	32	37
Total	**100**	**100**

Item contexts are spread across *personal*, *social* and *global* settings approximately in the ratio 1:2:1, and there is a roughly even selection of areas of application as listed in Figure 3.2.

The distributions for item types are given in Table 3.3.

Table 3.3
Approximate distribution of score points in science, by item types

	Percentage of score points	
Item types	**PISA 2006**	**PISA 2012**
Simple multiple-choice	35	32
Complex multiple-choice	27	34
Closed constructed-response	4	2
Open constructed-response	34	32
Total	**100**	**100**



REPORTING PROFICIENCY IN SCIENCE

PISA results are reported on a scale constructed using a generalised form of the Rasch model as described by Adams, Wilson and Wang (1997). For each domain (reading, mathematics and science), a scale is constructed with a mean score of 500 and standard deviation of 100 among OECD countries; accordingly, about two-thirds of students across OECD countries score between 400 and 600 points.

When science was the major assessment domain for the first time in 2006, six proficiency levels were defined on the science scale. These same proficiency levels will be used in the reporting of science results for PISA 2012. Proficiency at each of the six levels can be understood in relation to the kinds of scientific competencies that a student needs to attain at each level. Figure 3.10 presents a description of the scientific knowledge and skills which students possess at the various proficiency levels, with Level 6 being the highest level of proficiency. It also gives the level and scale score of each item belonging to the three units from the PISA 2006 assessment, which are used as examples throughout this chapter.

The capacity of students who performed below Level 1 in PISA 2006 (about 5.2% of students on average across OECD countries) could not be reliably described because not enough science items were located in this region of the scale. Level 2 was established as the baseline level of scientific literacy, defining the level of achievement on the PISA scale at which students begin to demonstrate the scientific knowledge and skills that will enable them to participate actively in life situations related to science and technology.

■ Figure 3.10 ■

Summary descriptions of the six proficiency levels in science

Level	Lower score limit	Examples of items at each level	What students can typically do at each level
6	707.9	ACID RAIN Q3 Full credit (717) GREENHOUSE Q3 (709)	At Level 6, students can consistently identify, explain and apply scientific knowledge and knowledge about science in a variety of complex life situations. They can link different information sources and explanations and use evidence from those sources to justify decisions. They clearly and consistently demonstrate advanced scientific thinking and reasoning, and they use their scientific understanding in support of solutions to unfamiliar scientific and technological situations. Students at this level can use scientific knowledge and develop arguments in support of recommendations and decisions that centre on personal, social or global situations.
5	633.3	GREENHOUSE Q2 Full credit (659)	At Level 5, students can identify the scientific components of many complex life situations, apply both scientific concepts and knowledge about science to these situations, and can compare, select and evaluate appropriate scientific evidence for responding to life situations. Students at this level can use well-developed inquiry abilities, link knowledge appropriately and bring critical insights to situations. They can construct explanations based on evidence and arguments based on their critical analysis.
4	558.7	PHYSICAL EXERCISE Q3 (583) GREENHOUSE Q2 Partial credit (568)	At Level 4, students can work effectively with situations and issues that may involve explicit phenomena requiring them to make inferences about the role of science or technology. They can select and integrate explanations from different disciplines of science or technology and link those explanations directly to aspects of life situations. Students at this level can reflect on their actions and they can communicate decisions using scientific knowledge and evidence.
3	484.1	PHYSICAL EXERCISE Q1 (545) GREENHOUSE Q1 (529) ACID RAIN Q3 Partial credit (513) ACID RAIN Q1 (506)	At Level 3, students can identify clearly described scientific issues in a range of contexts. They can select facts and knowledge to explain phenomena and apply simple models or inquiry strategies. Students at this level can interpret and use scientific concepts from different disciplines and can apply them directly. They can develop short statements using facts and make decisions based on scientific knowledge.
2	409.5	ACID RAIN Q2 (460)	At Level 2, students have adequate scientific knowledge to provide possible explanations in familiar contexts or draw conclusions based on simple investigations. They are capable of direct reasoning and making literal interpretations of the results of scientific inquiry or technological problem solving.
1	334.9	PHYSICAL EXERCISE Q2 (386)	At Level 1, students have such a limited scientific knowledge that it can only be applied to a few, familiar situations. They can present scientific explanations that are obvious and follow explicitly from given evidence.

Factors that determine difficulty of items assessing science achievement include:

- The general complexity of the context.
- The level of familiarity of the scientific ideas, processes and terminology involved.
- The length of the train of logic required to respond to a question – that is, the number of steps needed to arrive at an adequate response and the level of dependence of each step on the previous one.
- The degree to which abstract scientific ideas or concepts are required in forming a response.
- The level of reasoning, insight and generalisation involved in forming judgements, conclusions and explanations.

Question 3 of *GREENHOUSE* is an example of a difficult item, located at Level 6 on the PISA science scale. This question combines aspects of the two competencies, *identifying scientific issues* and *explaining phenomena scientifically*. As a first step to solving this problem, the student must be able to identify the change and measured variables and have sufficient understanding of the methods of investigation to recognise the influence of other factors. In addition, the student needs to recognise the scenario and identify its major components. This involves identifying a number of abstract concepts and their relationships in order to determine what "other" factors might affect the relationship between Earth's temperature and the amount of carbon dioxide emissions in the atmosphere. Thus, in order to respond correctly, a student must understand the need to control factors outside the changed and measured variables and must possess sufficient knowledge of "Earth systems" to identify at least one of the factors that should be controlled. Sufficient knowledge of "Earth systems" is considered the critical scientific skill involved, so this question is categorised as *explaining phenomena scientifically*.

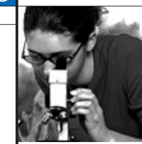
Question 1 of *PHYSICAL EXERCISE* is an example of an easy item, located at Level 1 on the PISA science scale below the baseline of scientific literacy. To gain credit, a student must correctly recall knowledge about the operation of muscles and formation of fat in the body, particularly the facts that when muscles are exercised they get an increased flow of blood and fats are not formed. This knowledge enables the student to accept the first statement of this complex multiple-choice question and reject the second one. In this item, no context needs to be analysed – the knowledge required has widespread currency and no relationships need investigating or establishing.

PISA 2006 results were also reported on three subscales corresponding to the three scientific competencies. These subscales used the same six proficiency levels as the combined scale, but with descriptors unique to each scale. In addition, country performance was compared on the bases of *knowledge about science* and the three main *knowledge of science* categories ("Physical systems", "Living systems", and "Earth and space systems").

While the analyses drawn from these kinds of comparisons could be valuable, caution should be used when relating performance to competencies and knowledge because the data come from classifying the same items in two ways that are not independent. All items classified as assessing the *identifying scientific issues* competency are *knowledge about science* items, and all *explaining phenomena scientifically* items are *knowledge of science* items (OECD, 2009, p. 44).

SUMMARY

The PISA definition of scientific literacy originates in the consideration of what 15-year-old students should know, value and be able to do as preparation for life in modern society. Central to the definition, and the science assessment, are the competencies that are characteristic of science and scientific enquiry: *identifying scientific issues, explaining phenomena scientifically,* and *using scientific evidence*. The ability of students to perform these competencies depends on their scientific knowledge, both knowledge of the natural world (i.e. knowledge of, chemistry, biology, Earth and space sciences, and technology) and *knowledge about science* itself (i.e. knowledge about "scientific enquiry" and "scientific explanations"), and their attitudes towards science-related issues.



This framework describes and illustrates the scientific competencies, knowledge and attitudes involved in the PISA definition of scientific literacy (see Figure 3.11), and outlines the format and structure of the PISA 2012 science assessment.

■ Figure 3.11 ■

Major components of the PISA definition of scientific literacy

Competencies	Knowledge	Attitudes
▪ Identifying scientific issues ▪ Explaining scientific phenomena ▪ Using scientific evidence	▪ Knowledge *of* science: – *Physical systems* – *Living systems* – *Earth and space systems* – *Technology systems* ▪ Knowledge *about* science – *Scientific enquiry* – *Scientific explanations*	▪ Interest in science ▪ Support for scientific enquiry ▪ Responsibility towards resources and environment

PISA science test items are grouped into units with each unit beginning with stimulus material that establishes the context for its items. The focus is on situations in which applications of science have particular value in improving the quality of life of individuals and communities. A combination of multiple-choice and constructed-response item types is used and some items involve partial credit scoring. Unlike PISA 2006, attitudinal items are not included in units in PISA 2012.

PISA 2012 science results will be reported on a single science scale having a mean of 500 and a standard deviation of 100, using the six levels of proficiency defined when science was the major assessment domain for the first time in 2006. Level 6 is the highest level of proficiency and Level 2 has been established as the baseline level of scientific literacy. Students achieving below Level 2 do not demonstrate the scientific knowledge and skills that will enable them to participate actively in life situations related to science and technology.

Notes

1. Throughout this framework, "natural world" includes the changes made by human activity, including the "material world" designed and shaped by technologies.

References

Adams, R.J., M. Wilson and **W.C. Wang** (1997), "The multidimensional random coefficients multinomial logit model", *Applied Psychological Measurement*, No. 21, pp. 1-23.

Blosser, P. (1984), *Attitude Research in Science Education*, ERIC Clearinghouse for Science, Mathematics and Environmental Education, Columbus, Ohio.

Bogner, F. and **M. Wiseman** (1999), "Toward measuring adolescent environmental perception", *European Psychologist*, No. 4, Vol. 3, pp. 139-151.

Bybee, R. (1997a), "Towards an understanding of scientific literacy", in W. Gräber and C. Bolte (eds.), *Scientific Literacy: An International Symposium,* Institute for Science Education at the University of Kiel (IPN).

Bybee, R. (1997b), *Achieving Scientific Literacy: From Purposes to Practices,* Heinemann, Portsmouth.

Bybee, R.W. and **B.J. McCrae** (eds.) (2009), *PISA Science 2006: Implications for Science Teachers and Teaching*, NSTA Press, Arlington, Virginia.

Eagles, P.F.J. and **R. Demare** (1999), "Factors influencing children's environmental attitudes", *The Journal of Environmental Education,* No. 30, Vol. 4, pp. 33-37.

Fensham, P.J. (1985), "Science for all: A reflective essay", *Journal of Curriculum Studies*, No. 17, Vol. 4, pp. 415-435.

Fensham, P.J. (2000), "Time to change drivers for scientific literacy", *Canadian Journal of Science, Mathematics, and Technology Education,* Vol. 2, pp. 9-24.

Fleming, R. (1989), "Literacy for a Technological Age", *Science Education,* No. 73, Vol. 4

Gardner, P.L. (1975), "Attitudes to science: A review", *Studies in Science Education,* No. 2, pp. 1-41.

Gardner, P.L. (1984), "Students' interest in science and technology: An international overview", in M. Lehrke, L. Hoffmann and P.L. Gardner (eds.), *Interests in Science and Technology Education* (pp. 15-34), Institute for Science Education at the University of Kiel (IPN).

Gauld, C. and **A.A. Hukins** (1980), "Scientific attitudes: A review", *Studies in Science Education,* No. 7, pp. 129-161.

Gräber, W and **C. Bolte** (eds.) (1997), *Scientific Literacy: An International Symposium*, Institute for Science Education at the University of Kiel (IPN).

Klopfer L.E. (1976), "A structure for the affective domain in relation to science education", *Science Education*, Vol. 60(3), pp. 299-312.

Koballa, T., A. Kemp and **R. Evans** (1997), "The spectrum of scientific literacy", *The Science Teacher,* No. 64, Vol. 7, pp. 27-31.

Kuhn, D. (1992), "Thinking as Argument", *Harvard Educational Review,* No. 62, Vol. 2.

LaForgia, J. (1988), "The affective domain related to science education and its evaluation", *Science Education,* Vol. 72, No. 4, pp. 407-421.

Law, N. (2002), "Scientific literacy: Charting the terrains of a multifaceted enterprise", *Canadian Journal of Science, Mathematics, and Technology Education,* No. 2, pp. 151-176.

Mayer, V.J. (ed.) (2002), *Global Science Literacy*, Kluwer Academic Publishers, Dordrecht.

Mayer, V.J. and **Y. Kumano** (2002), "The Philosophy of Science and Global Science Literacy", in V.J. Mayer (ed.), *Global Science Literacy,* Kluwer Academic Publishers, Dordrecht.

Millar, R. and **J. Osborne** (1998), *Beyond 2000: Science Education for the Future*, King's College London, School of Education, London.

Norris, S. and **L. Phillips** (2003), "How literacy in its fundamental sense is central to scientific literacy", *Science Education,* No. 87, Vol. 2.

OECD (2003), "Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)", Summary of the final report *Key Competencies for a Successful Life and a Well-Functioning Society,* OECD Publishing.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*, OECD Publishing.

OECD (2009), *PISA 2006 Technical Report*, OECD Publishing.

Osborne, J., S. Erduran, S. Simon and **M. Monk** (2001), "Enhancing the Quality of Argumentation in School Science", *School Science Review,* No. 82, Vol. 301.

Osborne, J., S. Simon and **S. Collins** (2003), "Attitudes towards science: a review of the literature and its implications", *International Journal of Science Education,* No. 25, Vol. 9, pp. 1049-1079.

Rickinson, M. (2001), “Learners and learning in environmental education: A critical review of the evidence”, *Environmental Education Research,* No. 7, Vol. 3, pp. 207-208.

Roberts, D. (1983), *Scientific Literacy: Towards Balance in Setting Goals for School Science Programs,* Science Council of Canada, Ottawa.

Schibeci, R.A. (1984), “Attitudes to science: An update“, *Studies in Science Education,* No. 11, pp. 26-59.

UNESCO (1993), *International Forum on Scientific and Technological Literacy for All: Final Report,* UNESCO, Paris.

Weaver, A. (2002), “Determinants of environmental attitudes: A five-country comparison“, *International Journal of Sociology,* No. 32, Vol. 1, pp. 77-108.

4

PISA 2012 Problem-Solving Framework

This chapter presents the framework underlying the PISA 2012 computer-based assessment of individual problem-solving competency, including the rationale for the assessment, the framework's research underpinnings and a definition of what is meant by problem-solving competency. The definition is discussed in detail, as are the three key domain elements of most importance for the assessment: the problem context, the nature of the problem situation, and the cognitive processes involved in solving a problem.

The general structure of the assessment and its computer delivery are described, including the test interface and the response formats employed. The distributions of items by problem nature and context, and according to cognitive process, are specified. The inclusion of problems that require the solver to interact with the problem situation to uncover necessary information not explicitly disclosed is highlighted. Sample items are presented with commentary, including an illustration of how response data (captured by the computer-delivery system) is used to enhance scoring.

INTRODUCTION

Problem-solving competency is a central objective within the educational programmes of many countries. The acquisition of increased levels of problem-solving competency provides a basis for future learning, for effective participation in society and for conducting personal activities. Citizens need to be able to apply what they have learnt to new situations. The study of individuals' problem-solving strengths provides a window into their capabilities to employ basic thinking and other general cognitive approaches to confronting challenges in life (Lesh and Zawojewski, 2007).

Background to the 2012 assessment

Problem solving was an additional assessment domain in PISA 2003. Some key findings of the survey were as follows (OECD, 2005):

- In some countries 70% of students could solve relatively complex problems, while in others less than 5% could do so.
- In most countries, more than 10% of students were unable to solve basic problems.
- On average in OECD countries, half of the students were unable to solve problems that are more difficult than basic problems.
- Patterns of within-country variation in students' problem-solving proficiency differed considerably across countries.
- Patterns of within-country differences between problem-solving proficiency and domain-related proficiencies (mathematics, reading and science) differed considerably across countries.

Since the 2003 problem solving assessment framework was developed (OECD, 2003a), considerable research has been carried out in the areas of complex problem solving, transfer, computer-based assessment of problem solving, and large-scale assessment of problem-solving competency (e.g. Blech and Funke, 2005; Funke and Frensch, 2007; Greiff and Funke, 2008; Klieme, 2004; Klieme et al., 2005; Leutner et al., 2004; Mayer, 2002; Mayer and Wittrock, 2006; O'Neil, 2002; Osman, 2010; Reeff et al., 2006; Wirth and Klieme, 2004). This research has led to advances in understanding and measuring individuals' problem-solving capabilities.

In addition, advances in software development tools and the use of networked computers have made greater efficiency possible and have increased the effectiveness of assessment, including the capability to administer dynamic and interactive problems, engage students' interest more fully and capture more information about the course of the problem-solving process. On this last point, computer delivery of assessment tasks makes it possible to record data about such things as the type, frequency, length and sequence of actions performed by students when responding to items.

It is appropriate, therefore, to once again make problem solving an assessment domain in PISA, but in doing so to devise a new framework and implement additional assessment methodologies that allow for the real-time capture of students' capabilities. In particular, the PISA 2012 assessment of problem solving is computer-based and the student's interaction with the problem is a central feature of the assessment.

PISA 2012 problem solving is an assessment of individual problem-solving competency. Collaborative problem-solving skills – the skills required to solve problems as a member of a group – are essential for successful future employment, where the individual is often a member of a team of diverse specialists working in separate locations. However, the significant measurement challenges associated with including collaborative tasks in a large-scale international survey such as PISA (Reeff et al., 2006), and the time needed to develop an appropriate computer-delivery platform, did not allow them to be a feature of the 2012 assessment.

A consistent research finding is that expert problem solving is dependent on domain-specific knowledge and strategies (e.g. Mayer, 1992; Funke and Frensch, 2007). The PISA 2012 assessment will avoid the need for expert knowledge as much as possible in order to focus on measuring the cognitive processes fundamental to problem solving. This also distinguishes the assessment from problem-solving tasks in the core PISA literacy domains of reading, mathematics and science, which call on expert knowledge in these areas.

Another conclusion that can be drawn from recent research is that authentic, relatively complex problems, particularly those that require direct interaction by the solver to uncover and discover relevant information, should be a central feature of the PISA 2012 problem-solving assessment. Examples are the problems commonly faced when using unfamiliar everyday devices such as remote controls, personal digital devices (e.g. mobile phones), home appliances and vending machines. Other examples arise in situations such as athletic training, animal husbandry, growing plants and social interactions. Problem-solving skills are necessary to achieve more than a basic level of skill when dealing with such situations and there is evidence that further skills, in addition to those involved in traditional reasoning-based problem solving, are required (e.g. Klieme, 2004). This is the first time that such "interactive problems" have been included in a large-scale international survey, which has been made possible by computer delivery of the assessment.

Problem-solving competency can be developed by high-quality education. Progressive teaching methods, like problem-based learning, inquiry-based learning, and individual and group project work, can be used to foster deep understanding and prepare students to apply their knowledge in novel situations. Good teaching promotes self-regulated learning and metacognition and develops the cognitive processes that underpin problem solving. It prepares students to reason effectively in unfamiliar situations, and to fill gaps in their knowledge by observation, exploration and interaction with unknown systems. The PISA 2012 computer-based assessment of problem solving aims to examine how students are prepared to meet unknown future challenges for which teaching of today's knowledge is not sufficient.

Problem solving in the OECD survey of adult skills

The OECD survey of adult skills is an assessment of reading component skills, literacy, numeracy and problem solving in technology-rich environments. It is a face-to-face sample household survey of people aged 16 - 65 years and was conducted for the first time in 2011, with results to be published in 2013.

The survey's assessment of "problem solving in technology-rich environments" differs from the PISA 2012 assessment of problem solving in two important aspects.[1] First, it is primarily concerned with "information-rich" problems. Examples include needing to locate and evaluate information on the Internet or on social networking sites, navigating through unfamiliar web pages and making decisions about what information is relevant and irrelevant for a task.

A second major difference is that problem solutions require the use of one or more computer software applications (file management, web browser, email and spreadsheet). In PISA, information and communication technology (ICT) is integral to the assessment of problem solving but it is not integral to its definition of problem solving. Only foundational ICT skills (based on use of a keyboard and mouse) are necessary to take the PISA computer-based assessment of problem solving. Software tools are common and are powerful aids for information-rich problem solving; a high level of ICT literacy is essential in this digital age. However, the PISA assessment focuses on the fundamental cognitive processes that are essential for successful problem solving with or without ICT assistance.

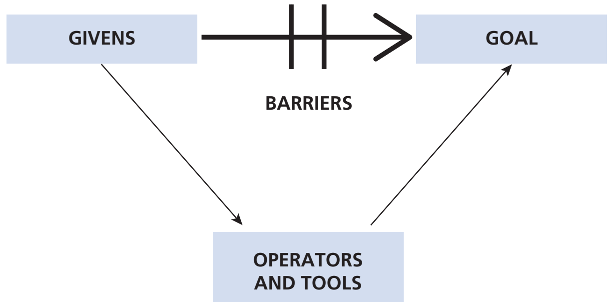
DEFINING PROBLEM-SOLVING COMPETENCY

The aim of the PISA 2012 problem-solving assessment is to assess individual problem-solving *competency*. Before defining what is meant by the term "problem-solving competency" in this context, it is important to clarify what is meant by the terms "problem" and "problem solving" by researchers in the field.

Definition of a problem

A problem exists when a person has a goal but doesn't know how to achieve it (Duncker, 1945). This definition is enlarged upon in Figure 4.1. The given state (givens) is the knowledge the person has about the problem at the outset and the operators are the admissible actions that can be performed to achieve the desired goal state (outcomes) with the assistance of the available tools. Barriers that must be overcome (e.g. lack of knowledge or obvious strategies) stand in the way of achieving the goal. Overcoming the barriers may involve not only cognition, but also motivational and affective factors (Funke, 2010).

■ Figure 4.1 ■
Problem situation



Source: Frensch and Funke, 1995.

As an example, consider the simple problem of finding the quickest route between two towns, when given a road map with estimated travel times marked and a calculator. The given state is the given information – the map with no route marked – and the goal state is the desired answer – the quickest route. The allowable actions (operators) are: selecting a possible route, calculating its total time and comparing it with the times for other routes. A tool (calculator) is available for assistance in adding times.

Definition of problem solving

Consistent with this understanding of what is meant by a problem, Mayer (1990) defines problem solving as cognitive processing directed at transforming a given situation into a goal situation when no obvious method of solution is available. This definition is widely accepted in the problem-solving community (e.g. see Klieme, 2004; Mayer and Wittrock, 2006; Reeff et al., 2006).

Definition of problem-solving competency

The PISA 2012 definition of problem-solving competency is grounded in these generally-accepted meanings of "problem" and "problem solving." It is as follows:

> *Problem-solving competency is an individual's capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious. It includes the willingness to engage with such situations in order to achieve one's potential as a constructive and reflective citizen.*

Not surprisingly, the first sentence of the definition is almost identical to the first part of the definition used for the PISA 2003 assessment of problem solving.[2] However, whereas the 2003 definition had only a cognitive dimension, with its latter part highlighting the cross-curricular nature of the assessment, an affective component has been introduced in the 2012 definition in line with the definition of competency as recognised by the OECD (OECD, 2003a).

What distinguishes the 2012 assessment of problem solving from the 2003 assessment is not so much the definition of problem-solving competency, but the mode of delivery of the 2012 assessment (computer-based) and the inclusion of problems that cannot be solved without the solver interacting with the problem situation.

In the following paragraphs, each part of the PISA 2012 definition of problem-solving competency is considered in turn to help clarify its meaning in relation to the assessment.

Problem-solving competency…

A competency involves far more than the basic reproduction of accumulated knowledge. It involves a mobilisation of cognitive and practical skills, creative abilities and other psychosocial resources such as attitudes, motivation and values (OECD, 2003b). The PISA 2012 assessment of problem-solving competency will not test simple reproduction of domain-based knowledge; rather it will focus on the cognitive skills required to solve unfamiliar problems that are encountered in life[3] and that lie outside traditional curricular domains.

Prior knowledge is important in solving problems. However, problem-solving competency involves the ability to acquire and use new knowledge, or to use old knowledge in a new way to solve novel problems (i.e. problems that are not routine).

…is an individual's capacity to engage in cognitive processing…

Problem solving occurs internally in an individual's cognitive system and can only be inferred indirectly by the person's actions and products. It involves representing and manipulating various types of knowledge in the problem solver's cognitive system (Mayer and Wittrock, 2006). Students' responses to assessment items – their exploration strategies, the representations they employ in modelling the problem, numerical and non-numerical answers, or extended explanations of how a problem was solved – will be used to make inferences about the cognitive processes they employed.

Creative (divergent) thinking and critical thinking are important components of problem-solving competency (Mayer, 1992). Creative thinking is a cognitive activity that results in finding solutions to a novel problem. Critical thinking accompanies creative thinking and is employed to evaluate possible solutions. The PISA 2012 assessment will target both components.

…to understand and resolve problem situations…

To what degree can individuals meet the challenges of a problem situation and move towards resolving it? In addition to explicit responses to items, the assessment aims to measure individuals' progress in solving a problem, including

the strategies they employ. Where appropriate these strategies are tracked by means of behavioural data captured by the computer-delivery system: the type, frequency, length and sequence of interactions made with the system can be captured and used in scoring or in subsequent analyses of student performance.

Problem solving begins with recognising that a problem situation exists and establishing an understanding of the nature of the situation. It requires the solver to identify the specific problem(s) to be solved and to plan and carry out a solution, along with monitoring and evaluating progress throughout the activity.

Often in real-world problems there may not be a unique or exact solution. In addition, the problem situation may change during the solving process, possibly due to interaction with the problem solver or as a result of its own dynamic nature. These complexities were addressed when setting assessment tasks with the aim of striking a balance between the authenticity of a situation and the practicality of the assessment.

…where a method of solution is not immediately obvious.…

The means of finding a solution path should not be immediately obvious to the problem solver. There will be barriers of various sorts in the way, or missing information. The assessment is concerned with non-routine problems and not routine ones (i.e. problems for which a previously learnt solution procedure is clearly applicable): the problem solver must actively explore and understand the problem and either devise a new strategy or apply a strategy learnt in a different context to work towards a solution.

The status of a problem – whether it is routine or not – depends on the solver's familiarity with the problem. A "problem" for one person may have an obvious solution to another person who is experienced and practised in solving such problems. Accordingly, care was taken to set problems that should be non-routine for the great majority of 15-year-olds.

It is not necessarily the case that the context or goals will themselves be unfamiliar to the solver; what is important is that the particular problems are novel or the ways of achieving the goals are not immediately obvious. The problem solver might need to explore or interact with the problem situation before attempting to solve the problem. Direct interaction is made feasible by the use of computer-delivered assessment in PISA 2012.

…It includes the willingness to engage with such situations…

Problem solving is personal and directed, that is, the problem solver's processing is guided by their personal goals (Mayer and Wittrock, 2006). The problem solver's individual knowledge and skills help determine the difficulty or ease with which obstacles to solutions can be overcome. However, the operation of such knowledge and skill is affected by motivational and affective factors such as beliefs (e.g. self-confidence) and feelings about one's interest and ability to solve the problem (Mayer, 1998).

In addition, the context of a problem (whether it is familiar and understood), the external resources available to the solver (such as access to tools), and the environment in which the solver operates (e.g. an examination setting) will affect the way a person approaches and engages with the problem.

Motivational and affective factors will not be measured in the problem-solving cognitive assessment but the student questionnaire will contain groups of items measuring perseverance and openness with respect to problem solving generally. In addition, the questionnaire will include some questions that gather information on students' problem-solving strategies (e.g. ask someone convenient, consult instructions, engage in unfocussed behaviour, give up) when confronted with specific problem situations.

…in order to achieve one's potential as a constructive and reflective citizen.

Competency is an important factor in the ways that individuals help to shape the world, not just cope with it: "…key competencies can benefit both individuals and societies" (Rychen and Salganik, 2003). Individuals should "manage their lives in meaningful and responsible ways by exercising control over their living and working conditions" (*ibid*). They need to be proficient problem solvers to achieve their potential as constructive, concerned and reflective citizens.

Scope of the assessment

The PISA 2012 problem-solving assessment will not include problems that require expert knowledge for their solution. In particular, problems that could reasonably be included in an assessment of one of the three PISA core domains will not be included. Assessment tasks will centre on everyday situations, with a wide range of contexts employed as a means of controlling for prior knowledge in general.

Mobilisation of prior knowledge is not sufficient to solve novel problems in many everyday situations. Instead of a straightforward application of previously mastered knowledge, existing knowledge needs to be re-organised and combined with new knowledge using a range of reasoning skills. Gaps in knowledge must be filled through observation and exploration of the problem situation. This often involves interaction with a new system to discover rules that in turn must be applied to solve the problem. Such problems are the major focus of the PISA 2012 problem-solving assessment, made possible by it being computer based.

ORGANISING THE DOMAIN

How the domain is represented and organised determines the assessment design and, ultimately, the evidence about student proficiencies that can be collected and reported. Many elements are part of the construct, not all of which can be taken into account and varied in an assessment such as PISA. The most important elements must be identified so that they can be varied to ensure construction of an assessment that contains items which have an appropriate range of difficulty and provide a broad coverage of the domain.

The domain elements of key importance for the PISA 2012 problem-solving assessment are as follows:

- The problem context: whether it involves a technological device or not, and whether the focus of the problem is *personal* or *social*.
- The nature of the problem situation: whether it is *interactive* or *static*.
- The problem-solving processes: the cognitive processes involved in solving a problem.

Items were developed to measure how well students perform when the various problem-solving processes are exercised within the two different types of problem situations across a range of contexts. Each of these key domain elements is discussed and illustrated in the following sections.

Problem context

An individual's familiarity and understanding of the problem context will affect how difficult the problem is to solve for that person. Two dimensions have been identified to ensure that assessment tasks sample across a range of contexts that are authentic and of interest to 15-year-olds: the setting (technology or not) and the focus (personal or social).

Problems set in a *technology* context have the functionality of a technological device as their basis. Examples include mobile phones, remote controls for appliances and ticket vending machines. Knowledge of the inner workings of these devices will not be required: typically, students are led to explore and understand the functionality of a device, as preparation for controlling the device or for troubleshooting its malfunctioning. Situations that give rise to other types of problems, such as route planning, task scheduling and decision making, have *non-technology* contexts.

Personal contexts include those relating primarily to the self, family and peer groups. *Social* contexts relate to situations typically encountered more broadly in the community or society in general (including at work and in undertaking further education). As illustrations, the context of an item about setting the time on a digital watch would be classified as *technology* and *personal*, whereas the context of an item requiring the construction of a basketball team roster would be classified as *non-technology* and *social*. The last section of this chapter gives further examples: the first sample unit described, which focuses on the rules that govern the functioning of an MP3 player, has a *technology* and *personal* context, and the second, about the seating plan for a birthday party, has a *non-technology* and *social* context.

Nature of the problem situation

How a problem is presented has important consequences for how it can be solved. Of crucial importance is whether the information about the problem disclosed to the problem solver at the outset is complete. This is the case for the quickest-route problem discussed earlier (see the section "Definition of a problem"). We refer to such problem situations as being *static*. The unit *BIRTHDAY PARTY*, described in the section "Illustrative PISA problem solving items" is an example of a *static* unit.

By contrast, problem situations may be *interactive*, meaning that exploration of the situation to uncover additional relevant information is possible.[4] Real-time navigation using a GPS system where traffic congestion is reported automatically or by query, presents such a situation. The unit *MP3 PLAYER*, described in the section "Illustrative PISA problem solving items" is *interactive*.

Interactive problem situations can be simulated in a test setting by means of a computer. Including interactive problem situations in the computer-based PISA 2012 problem-solving assessment allows a wide range of more authentic,

real-life scenarios to be presented than would otherwise be possible using pen-and-paper tests. Problems where the student explores and controls a simulated environment are a distinctive feature of the assessment.

A selection of static problem situations also is included in the assessment. The assessment of such problems has traditionally taken place using pen-and-paper tests. However, their computer-based assessment has many advantages including the capability of presenting a broader range of scenarios, involving multimedia elements such as animation; the availability of online tools; and, the use of a wide range of response formats that can be automatically coded.

Furthermore, some studies suggest that knowledge acquisition in exploring a problem in an interactive environment, and how that knowledge is applied, are competencies distinct from the typical skills used in solving static problems (see Klieme, 2004; Wirth and Klieme, 2004; Leutner and Wirth, 2005). Including a mixture of interactive and static problems in the PISA 2012 assessment therefore will provide a broader measure of problem-solving competency than has been possible with pen-and-paper instruments.

Interactive problem situations

Interactive problem situations often arise when encountering technological devices such as ticket vending machines, air-conditioning systems or mobile telephones for the first time, especially if the instructions for using them are not clear or are not available. Understanding how to control such devices is a problem faced universally in everyday life. In these situations it is often the case that some relevant information is not apparent at the outset. For example, the effect of applying an operation (say, pushing a button on a remote control) may not be known and cannot be deduced, but rather must be inferred by interacting with the scenario through actually performing the operation (pushing the button) and forming a hypothesis about its function based on the outcome. In general, some exploration or experimentation must be done to acquire the knowledge necessary to control the device. Another common scenario is when a person must troubleshoot a fault or malfunction in a device. Here a certain amount of experimentation must take place to collect data on the circumstances under which the device fails.

An interactive problem situation can be dynamic, meaning that its state might change of its own accord due to influences beyond the problem solver's control (i.e. without any intervention by the problem solver).[5] For example, in the case of a ticket-vending machine, if during a transaction no buttons are pressed for 20 seconds, the machine might reset. Such autonomous behaviour of a system must be observed and understood so it can be taken into account in attaining the desired goal (purchasing a ticket).

Static problem situations

Static problem situations can give rise to well-defined or ill-defined problems. In a well-defined problem, such as the quickest-route problem (see the section "Definition of a problem"), the given state, goal state and allowable operators are clearly specified (Mayer and Wittrock, 2006). The problem situation is not dynamic (i.e. does not change of its own accord during the course of solving the problem), all relevant information is disclosed at the outset and there is a single goal.

Other examples of well-defined problems are traditional logic puzzles such as the Tower of Hanoi and the water jars problems (see, for example, Robertson, 2001); decision-making problems, where the solver is required to understand a situation involving a number of well-defined alternatives and constraints so as to make a decision that satisfies the constraints (e.g. choosing the right pain killer given sufficient details about the patient, the complaint and the available pain killers); and, scheduling problems for projects such as building a house or producing computer software, where a list of tasks with durations and dependencies between tasks is given.

Mayer and Wittrock (2006) point out that "educational materials often emphasise well-defined problems, although most real problems are ill-defined [i.e. not well-defined]". These latter problems, which may be interactive or static in nature, often involve multiple goals which are in conflict so that progress towards one may detract from progress towards the other(s). Elaboration and weighing of priorities is required for the problem solver to achieve a balance between the goals (Blech and Funke, 2010). An example is finding the "best" route between two places – should this be the shortest route?, the likely quickest route?, the most straightforward route?, the route with minimum variation in time?, etc. A more complex example is designing a car where high efficiency, low cost, high safety and low environmental footprint may all be desired.

Problem-solving processes

Different authors conceive of the cognitive processes involved in solving a problem in different ways, but there is a great deal of commonality in their views. The processes identified are derived from the work on problem solving and reasoning of cognitive psychologists (e.g. Baxter and Glaser, 1997; Bransford et al., 1999; Mayer and Wittrock, 1996, 2006; Vosniadou

and Ortony, 1989), as well as by the seminal work of Polya (1945). Additionally, recent work on complex and dynamic problem solving (Blech and Funke, 2005, 2010; Funke and Frensch, 2007; Greiff and Funke, 2008; Klieme, 2004; Osman, 2010; Reeff et al., 2006; Wirth and Klieme, 2004) has been taken into account.

No assumption is made that the processes involved in solving a particular problem are sequential or that all of the processes listed are involved in solving a particular problem. As individuals confront, structure, represent and solve authentic problems representing emerging life demands, they may move to a solution in a way that transcends the boundaries of a linear, step-by-step model. Most of the information about the functioning of the human cognitive system now supports the view that it is capable of parallel information processing (Lesh and Zawojewski, 2007).

For the purposes of the PISA 2012 problem-solving assessment, the processes involved in problem solving are taken to be:

- *Exploring and understanding*
- *Representing and formulating*
- *Planning and executing*
- *Monitoring and reflecting*

Exploring and understanding. The objective here is to build mental representations of each of the pieces of information presented in the problem. This involves:

- exploring the problem situation: observing it, interacting with it, searching for information and finding limitations or obstacles; and
- understanding given information and information discovered while interacting with the problem situation; demonstrating understanding of relevant concepts.

Representing and formulating. The objective here is to build a coherent mental representation of the problem situation (i.e. a situation model or a problem model). To do this, relevant information must be selected, mentally organised and integrated with relevant prior knowledge. This may involve:

- representing the problem by constructing tabular, graphical, symbolic or verbal representations, and shifting between representational formats; and
- formulating hypotheses by identifying the relevant factors in the problem and their interrelationships; organising and critically evaluating information.

Planning and executing includes:

- planning, which consists of goal setting, including clarifying the overall goal, and setting sub-goals, where necessary; and devising a plan or strategy to reach the goal state, including the steps to be undertaken; and
- executing, which consists of carrying out a plan.

Monitoring and reflecting includes:

- monitoring progress towards the goal at each stage, including checking intermediate and final results, detecting unexpected events, and taking remedial action when required; and
- reflecting on solutions from different perspectives, critically evaluating assumptions and alternative solutions, identifying the need for additional information or clarification and communicating progress in a suitable manner.

Reasoning skills

Each of the problem-solving processes draws upon one or more reasoning skills. In understanding a problem situation, the problem solver may need to distinguish between facts and opinion; in formulating a solution, the problem solver may need to identify relationships between variables; in selecting a strategy, the problem solver may need to consider cause and effect; and, in communicating the results, the problem solver may need to organise information in a logical manner. The reasoning skills associated with these processes are embedded within problem solving. They are important in the PISA context since they can be taught and modelled in classroom instruction (e.g. Adey et al., 2007; Klauer and Phye, 2008).

Examples of reasoning skills employed in problem solving include deductive, inductive, quantitative, correlational, analogical, combinatorial and multidimensional reasoning. These reasoning skills are not mutually exclusive and often in practice problem-solvers move from one to another in gathering evidence and testing potential solution paths before

settling into the major use of one method over others in finding the solution to a given problem. Reasoning skills have been broadly sampled across the assessment items because the difficulty of an item is influenced by the complexity and types of reasoning involved in its solution.

ASSESSING PROBLEM-SOLVING COMPETENCY

Structure of the assessment

The duration of the PISA 2012 computer-delivered assessment is 40 minutes. A total of 80 minutes of problem-solving material is organised into four 20-minute clusters. Students from countries not participating in the optional computer-based assessment of mathematics and digital reading will do 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 will do two, one or none of the four problem-solving clusters according to a separate balanced rotation design.

As is normal for PISA assessments, items are grouped into units based around a common stimulus that describes the problem situation. To minimise the level of reading literacy required, stimulus material (and task statements) are as clear, simple and brief as possible. Animations, pictures or diagrams have been used to avoid lengthy passages of text. Numeracy demands have also been kept to a minimum with, for example, running totals provided where appropriate.

There are 16 units altogether, comprising a total of about 40 items that have an appropriate range of difficulty. This will enable the strengths and weaknesses of populations and key subgroups to be determined with respect to the cognitive processes involved in problem solving.

Functionality provided by computer delivery

A principal benefit of measuring problem-solving competency through a computer-based assessment is the opportunity to collect and analyse data that relate to processes and strategies, in addition to capturing and scoring intermediate and final results. This is likely to be a major contribution of the PISA 2012 assessment of problem solving. With appropriate item authoring, data such as the type, frequency, length and sequence of actions performed by students can be captured for this purpose.

Only foundational ICT skills are assumed in the assessment, such as keyboard use, using a mouse or touchpad, clicking radio buttons, drag-and-drop, scrolling, and use of pull-down menus and hyperlinks. Care will be taken to ensure that interference with the measurement of problem-solving competency by ICT demand and presentation is kept to a minimum.

Both units and items within units will be delivered in a fixed order, or "lockstep" fashion. The lockstep procedure means that students are not able to return to an item or unit once they have moved to the next one. Each time students click the *Next* button a dialog box will display a warning that they are about to move on to the next item and that it will not be possible to return to the previous item. At this point, students can either confirm they want to move on or cancel the action and return to the current item.

The appearance of the test interface is consistent across items (see Figure 4.2). For each item the stimulus material appears in the top part of the screen. The item question appears in the lower part of the screen, and is separated visually from the stimulus through the use of borders. The division of the screen into two parts varies from item to item so that scrolling is never required to see all information.

At the top right of the screen a timing bar appears which shows how much time is remaining in the assessment. Down the left edge of the screen another indicator of progress is given: the items in the test are listed in unit groups, with the current item number highlighted.

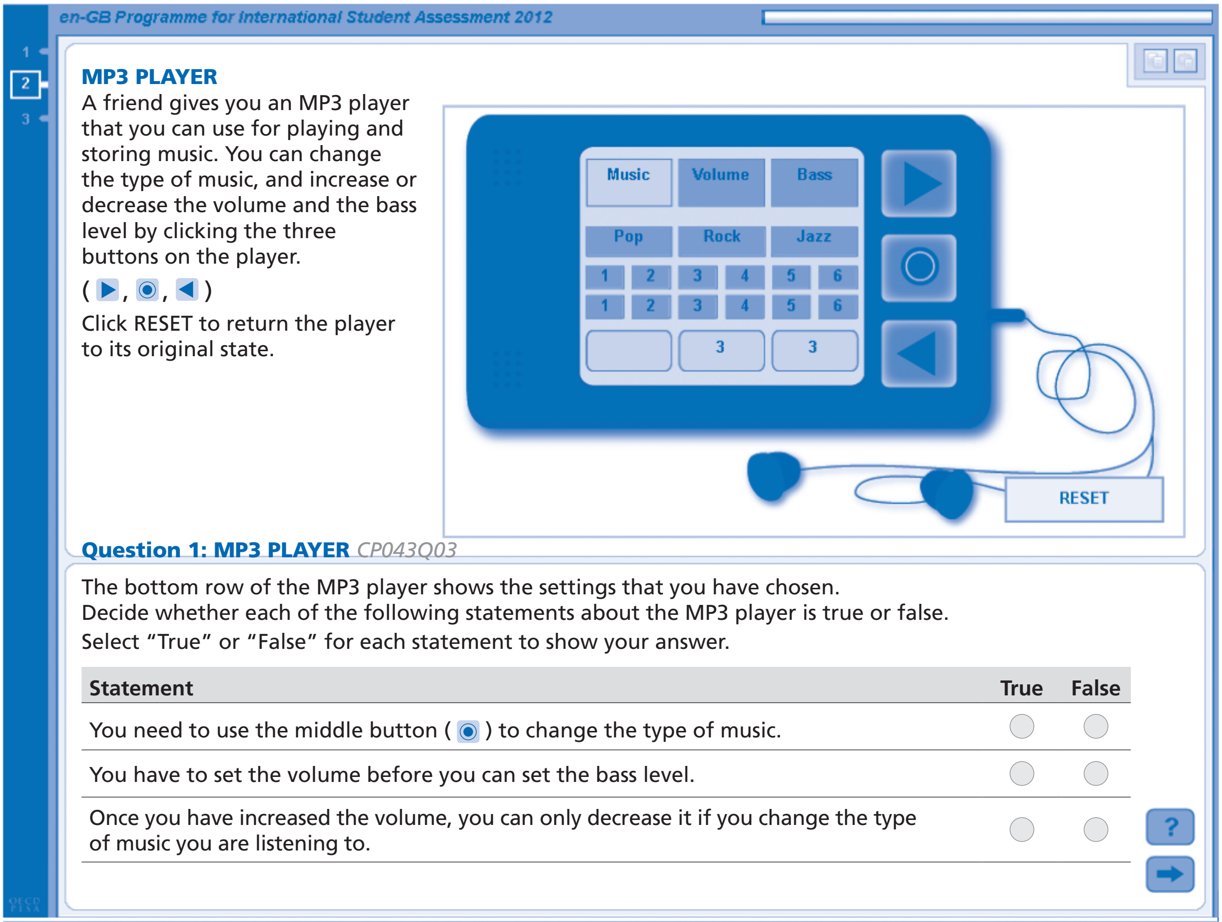
Task characteristics and difficulty

Generally, each item will focus on a single problem-solving process as far as possible. Accordingly, for some items, demonstrating a recognition of the problem will be sufficient; in others, describing a method of solution will be enough; in many, the actual solution(s) will be required with effectiveness and efficiency of method being important characteristics; in yet others, the task will be to evaluate proposed solutions and decide on the most appropriate solution for the problem posed. Including items that focus on one process is appropriate because although executing is often emphasised in classroom instruction, the major difficulties for most problem solvers involve representing, planning and self-regulating (Mayer, 2003).

Some problems are inherently more complex than others (Funke and Frensch, 2007). Furthermore, increased complexity generally means greater difficulty. Table 4.1 summarises task characteristics that are varied in the assessment to ensure that the items cover an appropriate range of difficulty. These characteristics are not mutually exclusive and can be regarded as forming four factors (Philpot et al., 2012) when the PISA 2012 problem set is analysed.

■ Figure 4.2 ■

The test interface



Response formats and coding

About one-third of the items require students to select their response(s) by clicking a radio button or by selecting from a drop-down menu. This includes simple multiple-choice items where there is one correct response to be selected, complex multiple-choice items where two or three separate multiple-choice selections must be made, and variations of these such as when there is more than one correct response to be selected from a list or multiple drop-down menus. All of these items are automatically coded.

Just over half the items require students to construct their responses but in a manner that can be automatically coded, such as by entering a number, dragging shapes, drawing lines between points or highlighting part of a diagram.

The remaining items require students to enter their responses in text boxes and need to be coded by experts. Such items are used in particular where it is considered important to ask students to explain their method or justify a selected response.

An online coding system has been developed to facilitate coding by experts. This eliminates the need for separate data entry, minimises the need for data cleaning, and allows coding to take place "off site" if desired.

The coding scheme for an item will allow for partial credit if appropriate, such as when multiple correct answers are required for full credit or when a correct strategy is employed but is not executed properly. Designated behaviours (such as exploration strategies) that provide reliable evidence about problem-solving competency over and above the fulfilment of task demands will be captured and will contribute to scoring.

Table 4.1
Task characteristics

Characteristic	Effect on task difficulty
Amount of information	The more information that has to be considered, the more difficult the task is likely to be.
Representation of information	Unfamiliar representations, and multiple representations (especially when information presented in different representations has to be related), tend to increase difficulty.
Degree of abstraction	How abstract or concrete the scenario is will affect the level of difficulty of a task. It is likely that the more abstract a scenario, the more difficult the task.
Familiarity of context	If a context is familiar to a problem solver, the solver may feel better equipped to tackle the problem.
Disclosure of information	The more that relevant information has to be discovered (e.g. effect of operations, autonomous behaviour, unanticipated obstacles), the more difficult the task is likely to be.
Internal complexity	The internal complexity of a task increases as the number of components or elements increases and they become more interrelated (due to dependencies or constraints). Tasks with a high level of internal complexity are likely to be harder than those with lower levels of complexity.
Distance to goal	The greater the number of steps needed to solve a problem, the more difficult it is likely to be.
Reasoning skills required	The difficulty of a task is influenced by the complexity and types of reasoning skills involved in its solution. Tasks that require the application of some types of reasoning (e.g. combinatorial reasoning) are likely to be harder than those that do not.

Interactive problems

Interactive problems can be built on underlying formal models whose parameters can be varied systematically to achieve differing degrees of difficulty. There are two well-used paradigms: linear difference equations and finite state machines.

With problem situations modelled by linear difference equations (also referred to as linear structural equations),[6] the problem solver must manipulate one or more input variables (such as controls for a climate control system) and consider the effect this has on one or more output variables (such as temperature and humidity); the output variables may also influence themselves so that the system is dynamic. Example contexts include remote controls, thermostats, paint mixing and ecosystems.

A finite state machine is a system with a finite number of states, input signals and output signals (Buchner and Funke, 1993).[7] The system's next state (and output signal) is uniquely determined by its current state and the specific input signal. With problem situations modelled by finite state machines, the problem solver must supply input signals (usually in the form of a sequence of button presses) to determine the effect on the system's states in an effort to understand its underlying structure and move it towards a goal state. Many everyday devices and contexts are governed or constrained by the rules of a finite state machine structure. Examples include digital watches, mobile phones, microwave ovens, MP3-players, ticket vending machines and washing machines.

The typical task demands for such interactive problems are as follows (see Blech and Funke (2005) and Greiff and Funke (2008) for additional details):

- Exploration: acquire knowledge of system structure either by active or directed exploration (interaction). *[Exploration strategies can be tracked and captured by the computer delivery system.]*
- Identification: give or complete a representation of the mental model of the system that is formed during exploration. This may be in drawing or text form. *[The accuracy of the model helps in assessing acquired causal knowledge.]*
- Control: practical application of acquired knowledge: transform a given state into a goal state and (for appropriate systems) maintain the goal state over time. A correct model of the system may be provided to minimise dependence on previous items. *[Transfer of acquired knowledge is assessed in this way.]*
- Explanation: describe strategies used to reach a goal; explain how a system works; or suggest causes of a malfunction of a device.

Students may already have some idea of the relationships between system variables in problem situations because of their familiarity with similar, actual devices. Such prior knowledge will vary between individuals and so a variety of common, everyday problem contexts will be used to help overcome this effect across the assessment. In addition, a

few more unusual but engaging game-like contexts will be included where the relationships must be inferred solely by manipulation and observation of system variables.

The difficulty of problems of these types is largely dependent on the internal complexity of the formal models underlying the situations. Problems of varying difficulty can be set by systematically varying this complexity, which is determined by the number of variables involved and how they are connected. For example, a problem involving only a few variables can be very easy if it only involves direct effects between input and output variables, but can be made extremely difficult by the inclusion of multiple effects and side effects between output variables.

Distribution of items

For the main survey, the percentage distribution of score points according to the cognitive processes involved in problem solving is given in Table 4.2. The ranges recommended by the Problem Solving Expert Group are included in parentheses. Highest weighting is given to *planning and executing* in recognition of the importance of being able to carry through a solution to a successful conclusion. Lower than average weight is given to *monitoring and reflecting* because it is an integral part of the other three processes and therefore also is assessed (indirectly) in items that target those processes.

Table 4.2
Approximate distribution of score points in problem solving, by process

Exploring and understanding	Representing and formulating	Planning and executing	Monitoring and reflecting	Total
21.4% *(20 - 25%)*	23.2% *(20 - 25%)*	41.1% *(35 - 45%)*	14.3% *(10 - 20%)*	100%

Table 4.3 indicates the percentage distribution of items across the two other key domain elements, problem context and nature of the problem situation. Once again, the recommended ranges are given in parentheses. The clear emphasis on interactive problems over static problems (a ratio of about 2:1) reflects the decision to concentrate on this important class of problems that, with the advantage of computer delivery, it is possible to include in a large-scale international survey for the first time. The greater emphasis on technology contexts over non-technology contexts recognises both the ever-increasing role played by technological devices in everyday life and their suitability for simulation in a computer-delivered test.

Table 4.3
Approximate distribution of items in problem solving, by problem nature and context*

	Technology context	Non-technology context	Total contexts
Static problem situation	11% *(10 - 15%)*	20% *(15 - 20%)*	31% *(25 - 35%)*
Interactive problem situation	45% *(40 - 45%)*	25% *(25 - 30%)*	70% *(65 - 75%)*
Total problem natures	55% *(50 - 60%)*	45% *(40 - 50%)*	100%

*Discrepancies in totals are due to rounding errors.

A roughly equal balance between personal and social contexts was recommended, subject to the constraints imposed in satisfying the recommended distributions for the key domain elements as expressed in Tables 4.2 and 4.3. The actual division of score points in the main survey is 59% personal and 41% social.

REPORTING PROBLEM-SOLVING COMPETENCY

Consistent with the other PISA domains, the results of the problem solving assessment will be summarised on a single, composite problem-solving scale with a mean of 500 and a standard deviation of 100.

Easier tasks appear at the lower end of the scale, and harder tasks appear at the upper end of the scale. In an attempt to capture and summarise the progression of task difficulty, the scale will be divided into levels.

Six levels of proficiency will be described to show how individuals' problem solving competency grows and develops, and to enable comparisons of student performance between and within participating countries and economies. There are not enough items in the survey to report on subscales.

Proficiency descriptions characterising typical student performance at each level will be developed by analysing the knowledge and skills required to answer the tasks at that level, and the characteristics of those tasks (see Table 4.1). It is expected that the following abilities will characterise high-performing students:

- Ability to plan and execute solutions that involve thinking a number of steps ahead and meeting mutiple constraints, to apply complex reasoning skills and to monitor progress towards a goal throughout the solution process, modifying plans where necessary.
- Ability to understand and make links across disparate pieces of information even when they are presented using unfamilar representations.
- Ability to interact with problems systematically and intenionally to discover undisclosed information.

Students who have not yet reached a baseline level of proficiency are expected to have at most, the following characteristics:

- Ability to plan and execute solutions that involve a small number of steps.
- Ability to solve problems involving one or two variables and no constraints, or only a single constraint.
- Ability to formulate simple rules, and discover undisclosed information when exploring in an unsystematic manner.

SUMMARY

The PISA 2012 assessment of problem solving is the second time that individual problem-solving competency has been assessed in PISA. In 2003, a pen-and-paper test of cross-disciplinary problem solving was part of the assessment. By contrast, the PISA 2012 assessment is computer-delivered, which enables the inclusion of items that require the solver to interact with the problem situation. Furthermore, problems that require disciplinary knowledge for their solution are avoided so as to focus on measuring the cognitive processes fundamental to problem solving.

For the purposes of PISA 2012, problem-solving competency is defined as an individual's capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious. It includes the willingness to engage with such situations in order to achieve one's potential as a constructive and reflective citizen.

The domain elements of key importance in developing assessment items are the problem context: *technological* or not, *personal* or *social*; the nature of the problem situation: interactive or static; and, the problem solving processes – the cognitive processes involved in problem solving: exploring and understanding, representing and formulating, planning and executing, monitoring and reflecting.

The nature of the problem situation is determined by whether the information disclosed to the solver at the outset about the problem situation is complete (static problems), or whether interaction with the problem situation is a necessary part of the solving activity in order to uncover additional information (interactive problems). Examples of interactive problems include problems commonly faced when using unfamiliar devices such as a new mobile phone or a ticket vending machine. This is the first time, made possible by computer delivery of the assessment, that such interactive problems have been included in a large-scale international survey.

Each test item, with its associated stimulus material, occupies a single computer screen and students proceed from item to item in a "lockstep" fashion. A variety of response formats are employed including selected-response and constructed-response that can be automatically coded (e.g. drag-and-drop), and free text entry that requires coding by experts. For some items, behavioural data that provides reliable evidence about problem-solving competency (e.g. exploration strategies) are captured and contribute to the scoring. Sample items are presented with commentary in the following section.

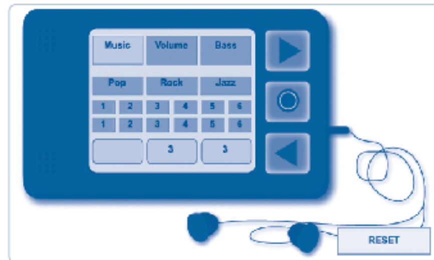
ILLUSTRATIVE PISA PROBLEM-SOLVING ITEMS

Items from two units that were included in the PISA 2012 field trial are described in this section.[8] For each unit a screenshot of the stimulus information is provided, together with a brief description of the context of the unit. This is followed by a screenshot and description of each item from that unit.

MP3 PLAYER

■ Figure 4.3 ■
MP3 player: Stimulus information

MP3 PLAYER
A friend gives you an MP3 player that you can use for playing and storing music. You can change the type of music, and increase or decrease the volume and the bass level by cliking the three buttons on the player.
(▶, ◉, ◀)
Click RESET to return the player to its original state.



In the unit MP3 Player, students are told that they have been given an MP3 Player by a friend. They do not know how it works and must interact it with it to find out, so the nature of the problem situation for each item in this unit is interactive. Since the focus of the unit is on discovering the rules that govern a device intended for use by an individual, the *context* of each item in the unit is *technology* and *personal*.

■ Figure 4.4 ■
MP3 player: Item 1

QUESTION 1

The bottom row of the MP3 player shows the settings that you have chosen.

Decide whether each of the following statements about the MP3 player is true or false. Select "True" or "False" for each statement to show your answer.

Statement	True	False
You need the use the middle button (◉) to change the type of music.	○	○
You have to set the volume before you can set the bass level.	○	○
Once you have increased the volume, you can only decrease it if you change the type of music you are listening to.	○	○

In the first item in the unit, students are given a series of statements about how the system works and asked to identify whether they are true or false. The statements offer scaffolding for students to explore the system. The problem-solving *process* for this item is *exploring and understanding*, and the exploration is guided but unrestricted. A "Reset" button is available which allows students at any time to return to the player to its initial state, and re-commence their exploration if desired. There is no restriction on the number of times this can be done. In the field trial this was a somewhat harder than average item, with 38% of students gaining full credit (True, False, False), due probably to the requirement that all three answers must be correct and the degree to which information has to be uncovered (no information is known about the system at the outset and so all knowledge of the rules of the system must come from interacting with it). Partial credit was not available for this item.

■ Figure 4.5 ■
MP3 player: Item 2

QUESTION 2

Set the MP3 player to Rock, Volume 4, Bass 2.

Do this using as few clicks as possible. There is no RESET button.

The second item in the unit is classified as *planning and executing*. In this item students must plan how to achieve, and then execute, a given goal. Of interest for this partial credit item is that process information captured by the computer-delivery system (in this case, how many steps the student takes to successfully reach the goal state) contributes to the score. The task is to be completed using as few clicks as possible and the option of returning the machine to its initial

state by pressing the "Reset" button is not available. If the number of clicks used (no more than 13) indicates that students have been efficient in reaching the goal they receive full credit, but if they reach the goal in a less efficient manner they only receive partial credit. The requirement for efficiency contributed to making it somewhat harder than average to obtain full credit for this item, though it was fairly easy to obtain partial credit. In the field trial, about 39% of students received full credit and about 33% received partial credit.

■ Figure 4.6 ■

MP3 player: Item 3

QUESTION 3

Shown below are four pictures of the MP3 player's screen. Three of the screens cannot happen if the MP3 player is working properly. The remaining screen shows the MP3 player when it is working properly.

Which screen shows the MP3 player working properly?

The third item in the unit is classified as *representing and formulating*, since it requires students to form a mental representation of the way the whole system works in order to identify which of four options given shows a state that is possible for this machine. The ability to return the player to its initial state which was present in the first, but absent in the second item of the unit, is again present here, so that the student may interact with the system as much or as little as they need to without restriction. Partial credit was not available for this item and in the field trial it was of similar difficulty to the first item in the unit, with 39% of students selecting the correct response (B).

■ Figure 4.7 ■

MP3 player: Item 4

QUESTION 4

Describe in the box below how you could change the way the MP3 player works so that there is no need to have the bottom button (◀).

You must still be able to change the type of music, and increase or decrease the volume and the bass level.

The final item in this unit is classified as *monitoring and reflecting*, and asks students to consider how the way the machine works might be reconceptualised. This item is one of a small number of constructed response items and requires expert scoring. Full credit answers are those that suggest how the MP3 player might still operate with only one button. There is no single correct answer, and students may think creatively in devising a solution, but the most obvious solution is to suggest changing the way the top button works so that once you reach the right side of the display, one more click takes you back to the left of the display. In the field trial, this was by far the hardest item in the unit, with only 25% of students gaining credit, no doubt due to the requirement for a constructed response and the item's degree of abstraction: students must imagine a hypothetical scenario and link it with their mental representation of how the system currently works, in order to describe a possible alternative functioning. Partial credit was not available for this item.

BIRTHDAY PARTY

■ Figure 4.8 ■
Birthday Party: Stimulus information

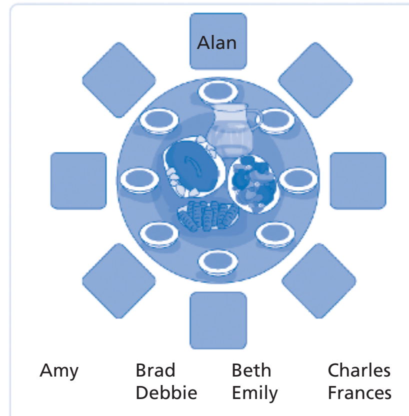
BIRTHDAY PARTY

It is Alan's birthday and he is having a party.

Seven other people will attend. Everyone will sit around the dining table.

The seating arrangement must meet the following conditions:

- Amy and Alan sit together
- Brad and Beth sit together
- Charles sits next to either Debbie or Emily
- Frances sits next to Debbie
- Amy and Alan do not sit next to either Brad or Beth
- Brad does not sit next to Charles or Frances
- Debbie and Emily do not sit next to each other
- Alan does not sit next to either Debbie or Emily
- Amy does not sit next to Charles



The scenario for this unit involves guests at a birthday party who must be placed around the dinner table in a way that satisfies nine specified conditions. The context for the unit is *non-technology* and *social*.

■ Figure 4.9 ■
Birthday Party: Item 1

QUESTION 1

Arrange the guests around the table to meet all of the conditions listed above. Use drag and drop to position the guests around the table.

In the only item in this unit students must drag and drop names to construct a seating plan subject to the nine conditions that are given. The item is therefore classified as *planning and executing*. Since all information that is necessary to solve the problem is given to students at the outset, the item is classified as *static*. Note that the item is *static* only in terms of the definition of the nature of the problem-solving situation. The response format (drag and drop) takes advantage of the capabilities of computer delivery: students can construct, review and revise their solution far more easily than would be possible in a paper-based version of this item. The item has partial credit scoring. For full credit, one of twelve possible solutions that meet all nine constraints must be found (e.g. Alan-Amy-Emily-Brad-Beth-Charles-Debbie-Frances); partial credit is given for solutions that meet only eight of the nine constraints (e.g. Alan-Amy-Emily-Brad-Beth- Debbie- Frances-Charles; here Charles does not satisfy the constraint of sitting next to Debbie or Emily). In the field trial, 54% of students gained at least partial credit on this item, with 43% gaining full credit. The difficulty of this item lies in the large number of conditions imposed, and the reasoning skills required to monitor and adjust partial solutions relative to these constraints until a complete solution is found.

Notes

1. The OECD survey of adult skills defines *problem solving in technology rich environments* as follows: “Problem solving in technology rich environments involves using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks” (PIAAC Expert Group in Problem Solving in Technology-Rich Environments, 2009, p.7).

2. “Problem solving is an individual’s capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricula areas that might be applicable are not within a single domain of mathematics, science or reading” (OECD, 2003a, p. 156).

3. Including those encountered in further education and work contexts.

4. The term “intransparent” is sometimes used to describe problems when complete information about the problem situation is not available at the outset (see Funke and Frensch, 1995).

5. The term “dynamic” is used by some researchers to describe any simulated physical system that a problem solver can interact with and receive feedback. In such cases, a problem situation that changes autonomously is sometimes termed “eigendynamic” (e.g. see Blech and Funke, 2005).

6. See Greiff and Funke (2008) who use the term “MicroDYN” to describe these systems. An earlier implementation of such a system is known as *Dynamis* – see Blech and Funke (2005).

7. Finite state machines for assessment purposes have been implemented under the name “MicroFin” – see *http://www.psychologie.uni-heidelberg.de/ae/allg_en/forschun/probleml.html.*

8. The two units “MP3 Player” and “Birthday Party” are available for viewing on the web at *http://cbasq.acer.edu.au*, using the credentials “public” and “access”. The interactive nature of MP3 Player can be best appreciated by trying it.

References

Adey, P., B. Csapó , A. Demetriou, J. Hautamäki and **M. Shayer** (2007), "Can we be intelligent about intelligence? Why education needs the concept of plastic general ability", *Educational Research Review* 2, pp. 75-97.

Baxter, G.P. and **R. Glaser** (1997), *An approach to analysing the cognitive complexity of science performance assessments (Technical Report 452)*, National Center for Research on Evaluation, Standards and Student Testing (CRESST), Los Angeles, California.

Blech, C. and **J. Funke** (2005), *Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology*, Deutsches Institut für Erwachsenenbildung, Bonn, *http://www.die-bonn.de/esprid/dokumente/doc-2005/blech05_01.pdf.*

Blech, C. and **J. Funke** (2010), "You cannot have your cake and eat it, too: How induced goal conflicts affect complex problem solving", *Open Psychology Journal 3*, pp. 42-53.

Bransford, J.D., A.L. Brown and **R.R. Cockling (eds.)** (1999), *How People Learn: Brain, Mind, Experience, and School*, National Academy Press, Washington, D.C.

Buchner, A. and **J. Funke** (1993), Finite-state automata: Dynamic task environments in problem-solving research, *The Quarterly Journal of Experimental Psychology*, Vol. 46A, No.1, pp. 83-118.

Duncker, K. (1945), "On problem solving", *Psychological Monographs*, Vol. 58, No. 3 (Whole No. 270).

Frensch, P.A. and **J. Funke,** (1995), "Definitions, traditions, and a general framework for understanding complex problem solving". In P. A. Frensch and J. Funke (eds.), *Complex problem solving: The European perspective* , Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 3-25.

Funke, J. (2010), "Complex problem solving: A case for complex cognition?", *Cognitive Processing*, Vol.11, pp. 133-142.

Funke, J. and **P.A. Frensch** (2007), "Complex problem solving: The European perspective – 10 years after", in D. H. Jonassen (ed.), *Learning to Solve Complex Scientific Problems*, Lawrence Erlbaum, New York, New York, pp. 25-47.

Greiff, S. and **J. Funke** (2008), *Indikatoren der Problemlöseleistung: Sinn und Unsinn verschiedener Berechnungsvorschriften. Bericht aus dem MicroDYN Projekt [Measuring Complex Problem Solving: The MicroDYN approach]*, Psychologisches Institut, Heidelberg.

Klauer, K. and **G. Phye** (2008), "Inductive reasoning: a training approach", *Review of Educational Research*, Vol. 78, No.1, pp. 85-123.

Klieme, E. (2004), "Assessment of cross-curricular problem-solving competencies", in J.H. Moskowitz and M. Stephens (eds.), *Comparing Learning Outcomes. International Assessments and Education Policy*, Routledge Falmer, London, pp. 81-107.

Klieme, E., D. Leutner and **J. Wirth (eds.)** (2005), *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA 2000 Studie*, VS Verlag für Sozialwissenschaften, Wiesbaden.

Lesh, R. and **J.S. Zawojewski** (2007), "Problem solving and modeling'', in F. Lester (ed.), *The Handbook of Research on Mathematics Teaching and Learning* (2nd ed.), National Council of Teachers of Mathematics, Reston, Virginia, and Information Age Publishing, Charlotte, North Carolina (joint publication), pp.763-804.

Leutner, D., E. Klieme, K. Meyer and **J. Wirth** (2004), "Problemlösen", in M. Prenzel et al. (PISA-Konsortium Deutschland) (eds.), PISA 2003: *Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*, Waxmann, Münster, pp.147-175.

Leutner, D. and **J. Wirth** (2005), "What we have learned from PISA so far: a German educational psychology point of view", *KEDI Journal of Educational Policy*, Vol. 2, No. 2, pp. 39-56.

Mayer, R.E. (1990), "Problem solving", in M. W. Eysenck (ed.), *The Blackwell Dictionary of Cognitive Psychology*, Basil Blackwell, Oxford, pp. 284-288.

Mayer, R.E. (1992), *Thinking, Problem solving, Cognition* (2nd ed.), Freeman, New York.

Mayer, R.E. (1998), "Cognitive, metacognitive, and motivational aspects of problem solving", *Instructional Science*, Vol. 26, pp. 49-63.

Mayer, R.E. (2002), "A taxonomy for computer-based assessment of problem solving", *Computers in Human Behavior*, Vol.18, pp. 623-632.

Mayer, R.E. (2003), *Learning and Instruction*, Merrill Prentice Hall, Upper Saddle River, New Jersey.

Mayer, R.E. and **M.C. Wittrock** (1996), Problem-solving transfer, in R. Calfee and R. Berliner (eds.), *Handbook of Educational Psychology*, Macmillan, New York, pp. 47-62.

Mayer, R.E. and **M.C. Wittrock** (2006), "Problem Solving", in P. A. Alexander and P. H. Winne (eds.), *Handbook of Educational Psychology* (2nd ed.), Lawrence Erlbaum Associates, Mahwah, New Jersey, Chapter 13.

OECD (2003a), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, PISA, OECD Publishing.

OECD (2003b), The definition and selection of competencies (DeSeCo): Executive summary of the final report, OECD Publishing, *www.oecd.org/dataoecd/47/61/35070367.pdf*.

OECD (2005), *Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competencies from PISA 2003*, PISA, OECD Publishing.

O'Neil, H.F. (2002), "Perspectives on computer-based assessment of problem solving", *Computers in Human Behavior*, Vol. 18, pp. 605-607.

Osman, M. (2010), "Controlling uncertainty: A review of human behavior in complex dynamic environments", *Psychological Bulletin*, Vol. 136, pp. 65-86.

Philpot, R., D. Ramalingam, J. Dossey and **B. McCrae** (2012), Paper presented at the 30th International Congress of Psychology, Cape Town, 22-27 July.

PIAAC Expert Group in Problem Solving in Technology-Rich Environments (2009), "PIAAC Problem Solving in Technology-Rich Environments: A Conceptual Framework", *OECD Education Working Papers*, No. 36, OECD Publishing.

Polya, G. (1945), *How to Solve It*, Princeton University Press, Princeton, New Jersey.

Reeff, J.-P., A. Zabal and **C. Blech** (2006), *The Assessment of Problem-Solving Competencies: A Draft Version of a General Framework*, Deutsches Institut für Erwachsenenbildung, Bonn, *http://www.die-bonn.de/esprid/dokumente/doc-2006/reeff06_01.pdf*, accessed 8 May 2008.

Robertson, S.I. (2001), *Problem Solving*, Psychology Press, East Sussex.

Rychen D.S. and **L.H. Salganik (eds.)** (2003), *Key Competencies for a Successful Life and a Well-Functioning Society*, Hogrefe and Huber, Göttingen.

Vosniadou, S. and **A. Ortony** (1989), *Similarity and Analogical Reasoning*, Cambridge University Press, New York.

Wirth, J. and **E. Klieme** (2004), "Computer-based assessment of problem solving competence", *Assessment in Education: Principles, Policy and Practice*, Vol. 10, No. 3, pp. 329-345.

5

PISA 2012 Financial Literacy Framework

PISA 2012 is the first large-scale international study to assess the financial literacy of young people. This framework is the first step in constructing a financial literacy assessment of international scope by providing an articulated plan for developing items, designing the instrument and providing a common language for discussion of financial literacy. This framework provides a working definition for financial literacy and organises the domain around the content, processes and contexts that are relevant for the assessment of 15-year-old students. Content areas described by the framework include *money and transactions, planning and managing finances, risk and reward* and *financial landscape*. The framework covers *identify financial information, analyse information in a financial context, evaluate financial issues, apply financial knowledge and understanding*, and the *education and work*, *home and family*, *individual* and *societal* contexts. These areas are illustrated with 10 items. Additionally, the framework discusses the relationship of financial literacy to non-cognitive skills and to both mathematics and reading literacy, and the measurement of students' financial behaviour and experience.

INTRODUCTION

The importance of financial literacy

In recent years, developed and emerging countries and economies have become increasingly concerned about the level of financial literacy of their citizens. This has stemmed in particular from shrinking public and private support systems, shifting demographic profiles including the ageing of the population, and wide-ranging developments in the financial marketplace. Concern was also heightened by the challenging economic and financial context with the recognition that lack of financial literacy was one of the factors contributing to ill-informed financial decisions and that these decisions could, in turn, have tremendous negative spill-overs (OECD INFE, 2009; OECD, 2009a; see also Gerardi et al., 2010, for empirical analysis of financial literacy and mortgage delinquency). As a result, financial literacy is now globally acknowledged as an important element of economic and financial stability and development; this is reflected in the recent G20 endorsement of the OECD/INFE (International Network on Financial Education) High-level Principles on National Strategies for Financial Education (G20, 2012; OECD INFE, 2012).

A series of tangible trends underpin the rising global interest in financial literacy as a key life skill. These are summarised below.

Risk shift

There has been a widespread transfer of risk from both governments and employers to individuals. Many governments are reducing or have reduced state-supported pensions, and some are reducing healthcare benefits. Defined-contribution pension plans are quickly replacing defined-benefit pension plans, shifting onto workers the responsibility to save for their own financial security after retirement. Traditional pay-as-you-go (PAYG) pension schemes are supplemented by new schemes in which the individual is subject to both revenue and investment risk. Most surveys show that a majority of workers are unaware of the risks they now have to face, and do not have sufficient knowledge or skill to manage such risks adequately, even if they are aware of them (OECD, 2008). Furthermore, the array of risks with financial implications is increasing: for example, individuals face the risks associated with longevity, credit, financial markets, and out-of-pocket healthcare.

Increased individual responsibility

The number of financial decisions that individuals have to make is increasing as a consequence of changes in the market and the economy. For instance, longer life expectancy means individuals need to ensure that they accumulate savings to cover much longer periods of retirement. People also need to assume more responsibility for funding personal or family healthcare needs. Moreover, increasing education costs make it important for parents to plan and invest adequately for their children's education. Even when individuals use the services of financial intermediaries and advisors, they need to understand what is being offered or advised. The individual is responsible for the financial product he or she decides to purchase, and the individual will face all the consequences of the choice. Individuals everywhere need to be financially literate to make informed and responsible decisions.

Increased supply of a wide range of financial products and services

In addition, in all countries, growing numbers of consumers have access to a wide range of financial products and services from a variety of providers and delivered through various channels. Improved levels of financial inclusion in emerging economies, developments in technology and deregulation have resulted in widening access to all kinds of financial products, from current accounts and remittances products to revolving credit and equity portfolios. The products available are also becoming more complex, and individuals are required to make comparisons across a number of factors such as the fees charged, interest rates paid or received, length of contract and exposure to risk. They must also identify appropriate providers and delivery channels from the vast array of possibilities, including community groups, traditional financial institutions, online banks and mobile phone companies.

Increased demand for financial products and services

Economic and technological developments have brought greater global connectedness and massive changes in communications and financial transactions, as well as in social interactions and consumer behaviour. Such changes have made it more important that individuals be able to interact with financial providers. In particular, consumers often need access to financial services (including banks and other providers such as post offices) in order to make and receive electronic payments like income, remittances and online transactions, as well as to conduct face-to-face transactions in societies where cash and cheques are no longer favoured. Those who cannot access such services often pay more for cash transactions, using informal financial services such as moneylenders or cheque cashers (see, for example, Kempson et al., 2005).

All of these trends have transferred the responsibility of major financial decisions to individuals. At the same time, they have both enlarged the options for the majority of the population (including new financial consumers) and increased the level of complexity they face. Against this backdrop, individuals are expected to be sufficiently financially literate to take the necessary steps to protect themselves and their relatives and ensure their financial well-being.

Expected benefits of financial education and improved levels of financial literacy

Existing empirical evidence shows that adults in both developed and emerging economies who have been exposed to financial education are subsequently more likely than others to save and plan for retirement (Bernheim et al., 2001; Cole et al., 2011; Lusardi, 2009). This evidence suggests a possible causal link between financial education and outcomes and indicates that improved levels of financial literacy can lead to positive behaviour change.

Other research, stemming largely from developed countries, and the United States in particular, indicates a number of potential benefits of being financially literate. There is mounting evidence that those with higher financial literacy are better able to manage their money, participate in the stock market and perform better on their portfolio choice, and that they are more likely to choose mutual funds with lower fees (Hastings and Tejeda-Ashton, 2008; Hilgert et al., 2003; Lusardi and Mitchell, 2008, 2011; Stango and Zinman, 2009; van Rooij et al., 2011; Yoong, 2011). Moreover, those who have greater financial knowledge are more likely to accumulate higher amounts of wealth (Lusardi and Mitchell, 2011).

Higher levels of financial literacy have been found to be related not only to asset building but also to debt and debt management, with more financially literate individuals opting for less costly mortgages and avoiding high interest payments and additional fees (Gerardi et al., 2010; Lusardi and Tufano, 2009a, 2009b; Moore, 2003).

In addition to the benefits identified for individuals, financial literacy is important to economic and financial stability for a number of reasons. Financially literate consumers can make more informed decisions and demand higher quality services, which will encourage competition and innovation in the market. They are also less likely to react to market conditions in unpredictable ways, less likely to make unfounded complaints and more likely to take appropriate steps to manage the risks transferred to them. All of these factors will lead to a more efficient financial services sector and potentially less costly financial regulatory and supervisory requirements. They can also ultimately help in reducing government aid (and taxation) aimed at assisting those who have taken unwise financial decisions – or no decision at all.

OECD activities in relation to financial education

In 2002, the OECD initiated a far-reaching financial education project to address governments' emerging concern about the potential consequences of low levels of financial literacy. This project is serviced by the OECD Committee on Financial Markets (CMF) and the Insurance and Private Pensions Committee (IPPC) in coordination with other relevant bodies including the Education Policy Committee. The project takes a holistic approach to financial-consumer issues that highlights how, alongside improved financial access, adequate consumer protection and regulatory frameworks, financial education has a complementary role to play in promoting the outcome of financial literacy.

One of the first milestones of the financial education project was the adoption of the *Recommendation on Principles and Good Practices for Financial Education and Awareness* by the OECD council (OECD, 2005a). Alongside these recommendations, the publication *Improving Financial Literacy: Analysis of Issues and Policies* details the reasons for focusing on financial education, and provides a first international overview of financial education work being undertaken in various countries (OECD, 2005b). This book also includes principles and good practices for policy makers and other stakeholders seeking to improve levels of financial literacy in their country. It is complemented by a global clearinghouse on financial education, the OECD International Gateway for Financial Education (*http://www.financial-education.org/home.html*), which gathers data, resources, research and news on financial education issues and programmes from around the world.

Recognising the increasingly global nature of financial literacy and education issues, in 2008 the OECD created the International Network on Financial Education (INFE) to benefit from and encompass the experience and expertise of developed and emerging economies. Currently more than 220 public institutions from more than 100 countries have joined the INFE. Members meet twice yearly to discuss the latest developments in their country and to collect evidence, develop analytical and comparative studies, methodologies, good practice, policy instruments and practical guidance on key priority areas. In this context, both financial education programmes in schools and the international measurement of financial literacy have been identified by the OECD and its INFE as top priority issues for which dedicated expert subgroups have been created to launch focused data collection and development work.

Financial education for youth and in schools

The focus on financial education for youth and in schools more specifically is not new. As mentioned, financial literacy is increasingly considered to be an essential life skill, and as early as 2005, the OECD Recommendation advised that "financial education should start at school. People should be educated about financial matters as early as possible in their lives" (OECD, 2005a). Two main reasons underpin this recommendation: the importance of focusing on youth, and the efficiency of providing financial education in schools. The OECD and its INFE have also developed guidelines for financial education in schools and guidance on learning frameworks on financial education which have been supported by the Ministers of Finance of the Asia-Pacific Economic Cooperation (APEC) in August 2012 and should be released by the end of 2012.[1]

Focus on youth

Younger generations are not only likely to face ever-increasing complexity in financial products, services and markets, but they are more likely to have to bear more financial risks in adulthood than their parents. In particular, they are likely to bear more responsibility for the planning of their own retirement savings and investments, and the coverage of their healthcare needs; and they will have to deal with more sophisticated and diverse financial products.

Because of the changes in the marketplace and social welfare systems (and particularly pension systems), current generations are unlikely to be able to learn from past generations. They will have to rely on their own knowledge or, given the complexities of new systems, make informed use of professional financial advice. Efforts to improve financial knowledge in the workplace or in other settings can be severely limited by a lack of early exposure to financial education and by a lack of awareness of the benefits of continuing financial education. It is therefore important to provide early opportunities for establishing the foundations of financial literacy.

In addition to preparing young people for their adult life, financial education in schools can also address the immediate financial issues facing young people. Children are often consumers of financial services from a young age. It is not uncommon for them to have accounts with access to online payment facilities or to use mobile phones (with various payment options) even before they become teenagers, and it is clear that financial literacy skills would be of benefit to them when using such products. Before leaving school, they may also face decisions about such issues as car insurance, savings products and overdrafts.

In many countries, at around the age of 15 to 18, young people (and their parents) face one of their most important financial decisions: that is, whether or not to invest in tertiary education. The gap in wages between college and non-college educated workers has widened in many economies. At the same time, the education costs borne by students and their families have increased, often leading to a reliance on credit (OECD, 2011). Figures published in March 2010 in the United Kingdom suggest that half of all British students expected to leave university owing over GBP 15 000 (Smithers, 2010).

It is important for people to be financially literate before they engage in major financial transactions and contracts. High-quality financial education programmes for young people can be essential in nurturing sound financial knowledge and behaviour in students from a young age, which they can draw on in the coming years (Ministerial Council for Education, Early Childhood Development and Youth Affairs [Australia], 2011).

Efficiency of providing financial education in schools

Research suggests that there is a link between financial literacy and family economic and educational background: those who are more financially literate disproportionately come from highly educated and families that hold a wide range of financial products (Lusardi et al., 2010). In order to provide equality of opportunity, it is important to offer financial education to those who would not otherwise have access to it. Schools are well positioned to advance financial literacy among all demographic groups and reduce financial literacy gaps and inequalities (including across generations).

Recognising both the importance of financial literacy for youth and the unique potential of school programmes to create more skilled and knowledgeable future generations, an increasing number of countries have embarked on the development of financial education programmes. These are either dedicated to youth generally or delivered through schools, and include those at national, regional and local levels as well as pilot exercises. A survey of individual financial literacy schemes supported by the European Commission (Habschick et al., 2007) found that most were directed at children and young people, and a broad stock-take exercise launched by the INFE subgroup on financial education in schools demonstrated that amongst the 32 countries/jurisdictions contributing to the survey, 21 had some programmes in schools (OECD, forthcoming).

The need for data

Policy makers, educators and researchers need high-quality data on levels of financial literacy in order to inform financial education strategies and the implementation of financial education programmes in schools by identifying priorities and measuring change across time.

Several countries have undertaken national surveys of financial literacy across their adult population and the OECD has recently piloted a questionnaire designed to capture levels of financial literacy amongst adults at an international level (Atkinson and Messy, 2012; OECD INFE, 2011). However, there are currently few data collection efforts on the levels of financial literacy amongst young people under the age of 18, and none that can be compared across countries. This is a serious omission as young people will soon be adults having to make ever more complex yet critical financial decisions, and the availability of data on their ability to address these challenges is essential to advance our knowledge on how well prepared the young are to face new and changing economic environments.

A robust measure of financial literacy amongst young people will provide information at a national level that can indicate whether the current approach to financial education is effective. In particular, it can help to identify issues that need addressing through schools or extra curricula activities or programmes that will enable young people to be properly and equitably equipped to make financial decisions in adulthood. It can also be used as a baseline from which to measure success and review school and other programmes in future years.

An international study provides additional benefits to policy makers and other stakeholders. Comparing levels of financial literacy across countries makes it possible to see which countries have the highest levels of financial literacy and begin to identify particularly effective national strategies and good practices. It will also be possible to recognise common challenges and explore the possibility of finding international solutions to the issues faced.

Against this backdrop, it is anticipated that the collection of robust and internationally comparable financial literacy data in the student population will provide policy makers, educators, curriculum and resource developers, researchers, and others with:

- information about gaps in financial knowledge amongst young people that can inform the development of more targeted programmes and policies;
- an indication of whether existing financial education in schools, where provided, is associated with higher levels of financial literacy;
- a means of comparing financial education strategies across countries;
- an opportunity to explore good practice by looking at the ranking of countries in terms of levels of financial literacy; and, ultimately,
- comparable data over time, allowing the assessment of the impact of financial education initiatives in schools and the identification of options for ongoing efficiency improvements.

There are other advantages to be gained from an international measurement exercise on financial literacy. In particular, the development of a financial literacy assessment framework that is applicable across countries provides national authorities with detailed guidance about the scope and operational definition of financial literacy without having to fund national studies. As noted in the article "Financial Literacy and Education Research Priorities", there is a gap in the research on financial literacy "related to the lack of consistency among researchers in how to define and measure programme success. There is a need for researchers to develop a clear understanding of what it means to be 'financially educated'"(Schuchardt et al., 2009).

The measurement of financial literacy in PISA

PISA 2012 is the first large-scale international study to assess the financial literacy of young people. PISA assesses the readiness of students for their life beyond compulsory schooling, and, in particular, their capacity to use knowledge and skills, by collecting and analysing cognitive and other information from 15-year-olds in many countries and economies. It is thus able to provide a rich set of comparative data that policy makers and other stakeholders can use to make evidence-based decisions. International comparative data on financial literacy can answer questions such as, "How well are young people prepared for the new financial systems that are becoming more global and more complex?" and "Who are the leaders in terms of financial literacy amongst youth?"

As with the core PISA domains of reading, mathematics and science, the main focus of the financial literacy assessment in PISA is on measuring the proficiency of 15-year-old students in demonstrating and applying knowledge and skills.

And like other PISA domains, financial literacy is assessed using an instrument designed to provide data that are valid, reliable and interpretable.

The first step in constructing an assessment that satisfies these three broad criteria is to develop an assessment framework. The main benefit of constructing an assessment framework is improved measurement, as it provides an articulated plan for developing the individual items and designing the instrument that will be used to assess the domain. A further benefit is that it provides a common language for discussion of the domain, and thereby increases understanding of what is being measured. It also promotes an analysis of the kinds of knowledge and skills associated with competency in the domain, thus providing the groundwork for building a described proficiency scale or scales that can be used to interpret the results.

The development of the PISA frameworks can be described as a sequence of the following six steps:

- development of a working definition for the domain and a description of the assumptions that underlie that definition;
- identification of a set of key characteristics that should be taken into account when constructing assessment tasks for international use;
- operationalisation of the set of key characteristics that will be used in test construction, with definitions based on existing literature and experience in conducting other large scale assessments;
- evaluation of how to organise the set of tasks constructed in order to report to policy makers and researchers on achievement in each assessment domain for 15-year-old students in participating countries;
- validation of the variables and assessment of the contribution each makes to understanding task difficulty across the various participating countries; and
- preparation of a described proficiency scale for the results.

DEFINING FINANCIAL LITERACY

In developing a working definition of financial literacy that can be used to lay down the groundwork for designing an international financial literacy assessment, the expert group looked both to existing PISA domain definitions of literacies, and to articulations of the nature of financial education.

PISA conceives of literacy as 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. PISA is forward looking, focusing on young people's ability to use their knowledge and skills to meet real-life challenges, rather than merely on the extent to which they have mastered specific curricular content (OECD, 2010a).

In its *Recommendation on Principles and Good Practices for Financial Education and Awareness*, the OECD defined financial education as "the process by which financial consumers/investors improve their understanding of financial products, concepts and risks and, through information, instruction and/or objective advice, develop the skills and confidence to become more aware of financial risks and opportunities, to make informed choices, to know where to go for help, and to take other effective actions to improve their financial well-being" (OECD, 2005a).

The FEG agreed that "understanding", "skills" and the notion of applying understanding and skills ("effective actions") were key elements of this definition. It was recognised, however, that the definition of financial education describes a process – education – rather than an outcome. What was required for the assessment framework was a definition encapsulating the outcome of that process in terms of competency or literacy.

The working definition of financial literacy for PISA 2012 is as follows:

> *Financial literacy is knowledge and understanding of financial concepts and risks, and the skills, motivation and confidence to apply such knowledge and understanding in order to make effective decisions across a range of financial contexts, to improve the financial well-being of individuals and society, and to enable participation in economic life.*

This definition, like other PISA domain definitions, has two parts. The first part refers to the kind of thinking and behaviour that characterises the domain. The second part refers to the purposes for developing the particular literacy.

In the following paragraphs, each part of the PISA 2012 definition of financial literacy is considered in turn to help clarify its meaning in relation to the assessment.

Financial literacy…

Literacy is viewed as an expanding set of knowledge, skills and strategies, which individuals build on throughout life, rather than as a fixed quantity, a line to be crossed, with illiteracy on one side and literacy on the other. Literacy involves more than the reproduction of accumulated knowledge, although measuring prior financial knowledge is an important element in the assessment. It also involves a mobilisation of cognitive and practical skills, and other resources such as attitudes, motivation and values. The PISA 2012 assessment of financial literacy draws on a range of knowledge and skills associated with development of the capacity to deal with the financial demands of everyday life in contemporary society.

…is knowledge and understanding of financial concepts and risks…

Financial literacy is thus contingent on some knowledge and understanding of fundamental elements of the financial world, including key financial concepts as well as the purpose and basic features of financial products. This also includes risks that may threaten financial well-being as well as insurance policies and pensions. It can be assumed that 15-year-olds are beginning to acquire this knowledge and gain experience of the financial environment that they and their families inhabit and the main risks they face. All of them are likely to have been shopping to buy household goods or personal items; some will have taken part in family discussions about money and whether what is wanted is actually needed or affordable; and a sizeable proportion of them will have already begun to earn and save money. Some students already have experience of financial products and commitments through a bank account or a mobile phone contract. A grasp of concepts such as interest, inflation, and value for money are soon going to be, if they are not already, important for their financial well-being.

…and the skills,…

These skills include such generic cognitive processes as accessing information, comparing and contrasting, extrapolating and evaluating – applied in a financial context. They include basic skills in mathematical literacy such as the ability to calculate a percentage or to convert from one currency to another, and language skills such as the capacity to read and interpret advertising and contractual texts.

…motivation and confidence…

Financial literacy involves not only the knowledge, understanding and skills to deal with financial issues, but also non-cognitive attributes: the motivation to seek information and advice in order to engage in financial activities, the confidence to do so and the ability to manage emotional and psychological factors that influence financial decision making. These attributes are considered as a goal of financial education, as well as being instrumental in building financial knowledge and skills.

…to apply such knowledge and understanding in order to make effective decisions…

PISA focuses on the ability to activate and apply knowledge and understanding in real-life situations rather than the reproduction of knowledge. In assessing financial literacy, this translates into a measure of young people's ability to transfer and apply what they have learnt about personal finance into effective decision-making. The term "effective decisions" refers to informed and responsible decisions that satisfy a given need.

…across a range of financial contexts…

Effective financial decisions apply to a range of financial contexts that relate to young people's present daily life and experience, but also to steps they are likely to take in the near future as adults. For example, young people may currently make relatively simple decisions such as how they will use their pocket money or, at most, which mobile phone contract they will choose; but they may soon be faced with major decisions about education and work options with long-term financial consequences.

…to improve the financial well-being of individuals and society…

Financial literacy in PISA is primarily conceived of as personal financial literacy, distinguished from economic literacy, which includes both broader concepts such as the theories of demand and supply, market structures and so on. Financial literacy is concerned with the way individuals understand, manage and plan their own and their households' – which often means their families' – financial affairs. It is recognised, however, that good understanding, management and planning on the part of individuals has some collective impact on the wider society, in contributing to national and even global stability, productivity and development.

…and to enable participation in economic life.

Like the other PISA literacy definitions, the definition of financial literacy implies the importance of the individual's role as a thoughtful and engaged member of society. Individuals with a high level of financial literacy are better equipped to make decisions that are of benefit to themselves, and also to constructively support and critique the economic world in which they live.

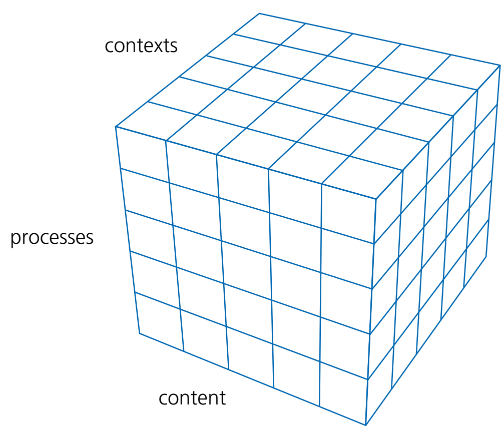
ORGANISING THE DOMAIN

How the domain is represented and organised determines the assessment design, including item development and, ultimately, the evidence about student proficiencies that can be collected and reported. Many elements are part of the concept of financial literacy, not all of which can be taken into account and varied in an assessment such as PISA. It is necessary to select the elements that will best ensure construction of an assessment comprising tasks with an appropriate range of difficulty and a broad coverage of the domain.

A review of approaches and rationales adopted in previous large-scale studies, and particularly in PISA, shows that most consider the relevant content, processes and contexts for assessment as they specify what they wish to assess. Content, processes and contexts can be thought of as three different perspectives on the area to be assessed, as shown in Figure 5.1.

■ Figure 5.1 ■

A model for organising the domain for an assessment framework



Content comprises the areas of knowledge and understanding that are essential in the area of literacy in question.

Processes describes the mental strategies or approaches that are called upon to negotiate the material.

Contexts refers to the situations in which the domain knowledge, skills and understandings are applied, ranging from the personal to the global.

The steps of identifying and weighting the different elements or categories within each perspective, and then ensuring that the set of tasks in the assessment adequately reflects these categories, are used to ensure coverage and validity of the assessment. The three perspectives are also helpful in thinking about how achievement in the area is to be reported.

The following section presents a discussion of each of the three perspectives and the framework categories into which they are divided. The section includes the types of tasks that a student may be asked to complete. Examples of items drawn from the PISA 2012 field trial are included to illustrate the perspectives and categories. While they are representative of those used in the main survey, these particular items are not used in the assessment instrument: only secure, unpublished items are used for this purpose, to protect the integrity of the data that is collected to measure student proficiency.

Content

The content of financial literacy is conceived of as the areas of knowledge and understanding that must be drawn upon in order to perform a particular task. A review of the content of existing financial literacy learning frameworks from a wide range of countries (Australia, Brazil, England, Japan, Malaysia, the Netherlands, New Zealand, Northern Ireland, Scotland, South Africa and the United States) indicated that there is some consensus on the financial literacy

content areas (OECD, forthcoming). The data analysis notably showed that the content of financial education in schools was – albeit with cultural differences – relatively similar, and that it was possible to identify a series of topics commonly included in these frameworks. These form the four content areas for PISA financial literacy: *money and transactions, planning and managing finances, risk and reward,* and *financial landscape*.

Money and transactions

This content area includes the awareness of the different forms and purposes of money and handling simple monetary transactions such as everyday payments, spending, value for money, bank cards, cheques, bank accounts and currencies. Tasks in this content area can, for example, ask students to show that they:

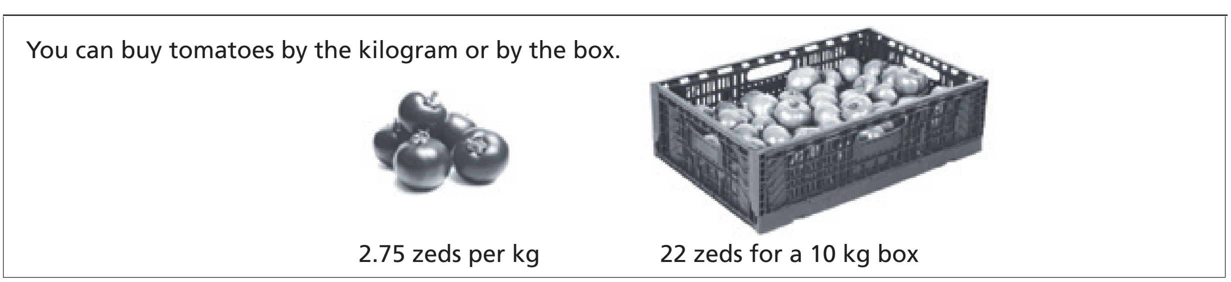
- Are aware of the different forms and purposes of money:
  - recognise bank notes and coins;
  - understand that money is used to exchange goods and services;
  - can identify different ways to pay for items, in person or via the Internet;
  - recognise that there are various ways of receiving money from other people and transferring money between people or organisations; and
  - understand that money can be borrowed or lent, and the reasons for paying or receiving interest.
- Are confident and capable at handling and monitoring transactions:
  - can use cash, cards and other payment methods to purchase items;
  - can use cash machines to withdraw cash or to get an account balance;
  - can calculate correct change;
  - can work out which of two consumer items of different sizes would give better value for money, taking into account the individual's specific needs and circumstances; and
  - can check transactions listed on a bank statement and note any irregularities.

The following example from the unit *AT THE MARKET* illustrates a task that requires students to apply the concept of value for money. In this question, and in many others, the unit of currency is the imaginary Zed. PISA questions often refer to situations that take place in the fictional country of Zedland, where the Zed is the unit of currency. This artifice (about which students are informed at the beginning of the testing session) has been introduced to enhance comparability across countries.

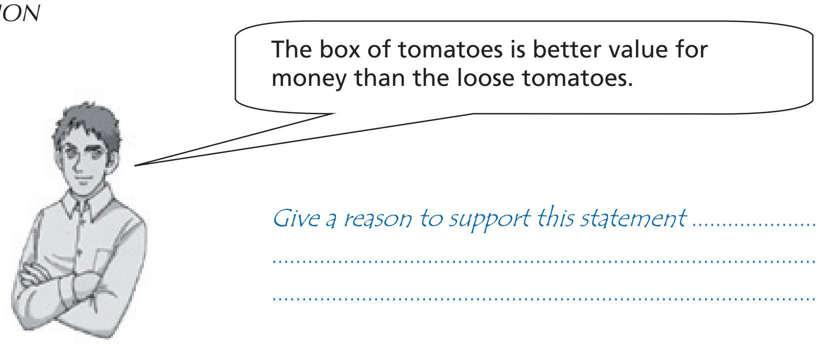
Illustrative PISA Financial literacy item 1 – AT THE MARKET

■ Figure 5.2 ■

Item for the unit *AT THE MARKET*



QUESTION



Give a reason to support this statement ....

....

....

Using an everyday context, shopping for groceries, this item addresses the basic concept of value for money. Questions about the buying of goods are generally categorised as being in the content area of money and transactions. To gain credit for this item, students have to show they have compared the two ways of buying tomatoes using a common point of comparison. Examples of answers that would receive credit are:

- It is 2.75 zeds/kg for loose tomatoes but only 2.2 zeds/kg for boxed tomatoes.
- Because 10 kg of loose tomatoes would cost 27.50 zeds.
- You get more tomato for each zed you spend when buying the box.

In the field trial, three-quarters of all students were able to analyse the information and explain that the price per kilogram of the boxed tomatoes is less than the price per kilogram of the loose tomatoes.

Planning and managing finances

Income and wealth need planning and managing over both the short term and long term. This content area includes:

- Knowledge and ability to monitor income and expenses:
  - identify various types of income and measures of income (e.g. allowances, salary, commission, benefits, hourly wage, and gross and net income); and
  - draw up a budget to plan regular spending and saving.
- Knowledge and ability to make use of income and other available resources in the short and long terms to enhance financial well-being:
  - understand how to manipulate various elements of a budget, such as identifying priorities if income does not meet planned expenses, or finding options for reducing expenses or increasing income in order to increase levels of savings;
  - assess the impact of different spending plans and be able to set spending priorities in the short and long term;
  - plan ahead to pay future expenses: for example, working out how much needs to be saved each month to make a particular purchase;
  - understand the purposes of accessing credit and the ways in which expenditure can be smoothed over time through borrowing or saving;
  - understand the idea of building wealth, the impact of compound interest on savings, and the pros and cons of investment products;
  - understand the benefits of saving for long term goals or anticipated changes in circumstance (such as living independently); and
  - understand how government taxes and benefits impact on planning and managing finances.

The example *SPENDING CHOICES* presented below illustrates an item addressing *planning and managing finances* in a context that is relevant to 15-year-olds as they think about their lives in the near future.

Illustrative PISA Financial literacy item 2 – SPENDING CHOICES

■ Figure 5.3 ■

Item for the unit *SPENDING CHOICES*

Claire and her friends are renting a house.
They have all been working for two months.
They do not have any savings.
They are paid monthly and have just received their wages.
They have made this "To do" list.

To do
- Get cable TV
- Pay the rent
- Buy outdoor furniture

QUESTION

Which of the tasks on the list are likely to need prompt attention from Claire and her friends?

Circle "Yes" or "No" for each task.

Task	Is the task likely to need prompt attention?
Get cable TV	Yes / No
Pay the rent	Yes / No
Buy outdoor furniture	Yes / No

This question in *SPENDING CHOICES* asks students to evaluate spending priorities for the home when working within a budget, distinguishing between wants and needs. It therefore falls under the content area of *planning and managing finances*. Over three-quarters of all students in the field trial scored full credit on this item by circling "No", "Yes" and "No" in that order, thus identifying that, of the three tasks, only paying the rent requires prompt attention for Claire and her housemates.

Another instance from the *planning and managing finances* content category is provided later, in the example *TRAVEL MONEY* which asks students to plan spending and saving in order to pay a future expense.

Risk and reward

Risk and reward is a key area of financial literacy, incorporating the ability to identify ways of managing, balancing and covering risks and an understanding of the potential for financial gains or losses across a range of financial contexts. There are two types of risk of particular importance in this domain. The first relates to financial losses that an individual cannot bear, such as those caused by catastrophic or repeated costs. The second is the risk inherent in financial products, such as credit agreements with variable interest rates, or investment products.

This content category includes:

- Recognising that certain financial products (including insurance) and processes (such as saving) can be used to manage and offset various risks (depending on different needs and circumstances):
  - knowing how to assess whether insurance may be of benefit.
- Applying knowledge of the ways to manage risk including the benefits of diversification and the dangers of default on payment of bills and credit agreements to decisions about:
  - limiting the risk to personal capital;
  - various types of investment and savings vehicles, including formal financial products and insurance products, where relevant; and
  - various forms of credit, including informal and formal credit, unsecured and secured, rotating and fixed term, and those with fixed or variable interest rates.
- Knowing about and managing risks and rewards associated with life events, the economy and other external factors, such as the potential impact of:
  - theft or loss of personal items, job loss, birth or adoption of a child, deteriorating health;
  - fluctuations in interest rates and exchange rates; and
  - other market changes.
- Knowing about the risks and rewards associated with substitutes for financial products; in particular:
  - saving in cash, or buying property, livestock or gold; and
  - borrowing from informal lenders.

An illustration from the *risk and reward* content category is provided in the example *MOTORBIKE INSURANCE*.

Illustrative PISA Financial literacy item 3 – MOTORBIKE INSURANCE

▪ Figure 5.4 ▪

Item for the unit *MOTORBIKE INSURANCE*

Last year, Steve's motorbike was insured with the PINSURA insurance company.
The insurance policy covered damage to the motorbike from accidents and theft of the motorbike.

QUESTION

Steve plans to renew his insurance with PINSURA this year, but a number of factors in Steve's life have changed since last year.

How is each of the factors in the table likely to affect the cost of Steve's motorbike insurance this year?

Circle "Increases cost", "Reduces cost" or "Has no effect on cost" for each factor.

Factor	How is the factor likely to affect the cost of Steve's insurance?
Steve replaced his old motorbike with a much more powerful motorbike	Increases cost / Reduces cost / Has no effect on cost
Steve has painted his motorbike a different colour	Increases cost / Reduces cost / Has no effect on cost
Steve was responsible for two road accidents last year	Increases cost / Reduces cost / Has no effect on cost

Motorbike insurance falls under the content area of *risk and reward* because insurance is a product designed specifically to protect individuals against risks and financial losses that they would not otherwise be able to bear. The question relies on students understanding that the higher their risk exposure is with regards to measurable criteria, the more it will cost them to buy appropriate insurance. Half of all students scored full credit on this question in the field trial by recognising that the first and third factors increase the cost of the insurance while the second factor has no effect.

Another illustration of the *risk and reward* content category is provided in the example *SHARES*, which requires students to be familiar with how a potentially risky product works.

Financial landscape

This content area relates to the character and features of the financial world. It covers knowing the rights and responsibilities of consumers in the financial marketplace and within the general financial environment, and the main implications of financial contracts. Information resources and legal regulation are also topics relevant to this content area. In its broadest sense, *financial landscape* also incorporates an understanding of the consequences of changes in economic conditions and public policies, such as changes in interest rates, inflation, taxation or welfare benefits. Tasks associated with this content area include:

- Knowledge of rights and responsibilities, and ability to apply it:
  - understand that buyers and sellers have rights, such as being able to apply for redress;
  - understand that buyers and sellers have responsibilities, such as:
    - consumers/investors giving accurate information when applying for financial products;
    - providers disclosing all material facts; and
    - consumers/investors being aware of the implications of one of the parties not doing so.
  - recognise the importance of the legal documentation provided when purchasing financial products or services and the importance of understanding the content.
- Knowledge and understanding of the financial environment, including:
  - identifying which providers are trustworthy, and which products and services are protected through regulation or consumer protection laws;
  - identifying whom to ask for advice when choosing financial products, and where to go for help in relation to financial matters; and
  - awareness of financial crimes such as identity theft and scams, and knowledge of how to take appropriate precautions.

- Knowledge and understanding of the impact of financial decisions including on others:
  - understand that individuals have choices in spending and saving and each action can have consequences for the individual and for society; and
  - recognise how personal financial habits, actions and decisions impact at individual, community, national and international level.
- Knowledge of the influence of economic and external factors:
  - aware of the economic climate and understand the impact of policy changes such as reforms related to the funding of post-school training;
  - understand how the ability to build wealth or access credit depends on economic factors such as interest rates, inflation and credit scores; and
  - understand that a range of external factors, such as advertising and peer pressure, can affect individuals' financial choices.

An illustration of an item that reflects financial landscape by focusing on financial crimes is presented in the example *BANK ERROR*.

Illustrative PISA Financial literacy item 4 – BANK ERROR

■ Figure 5.5 ■

Item for the unit *BANK ERROR*

David banks with ZedBank. He receives this e-mail message.

Dear ZedBank member,

There has been an error on the ZedBank server and your Internet login details have been lost.

As a result, you have no access to Internet banking.

Most importantly your account is no longer secure.

Please click on the link below and follow the instructions to restore access.
You will be asked to provide your Internet banking details.

https://ZedBank.com

QUESTION

Which of these statements would be good advice for David?

Circle "Yes" or "No" for each statement.

Statement	Is this statement good advice for David?
Reply to the e-mail message and provide his Internet banking details	Yes / No
Contact his bank to inquire about the e-mail message	Yes / No
If the link is the same as his bank's website address, click on the link and follow the instructions	Yes / No

Internet banking is part of the broader *financial landscape* in which students are likely to participate, either now or in the near future. In this environment they may be exposed to financial fraud. *BANK ERROR* investigates whether they know how to take appropriate precautions. In this question, students are asked to respond appropriately to a financial scam e-mail message. They must evaluate the presented options and recognise that the second piece of advice is the only one that can be considered good advice. In the field trial, just over 40% of the students gained full credit for this item, by responding "No", "Yes", "No" in that order.

Processes

The process categories relate to cognitive processes. They are used to describe students' ability to recognise and apply concepts relevant to the domain, and to understand, analyse, reason about, evaluate and suggest solutions. In PISA financial literacy, four process categories have been defined: *identify financial information, analyse information in a financial context, evaluate financial issues* and *apply financial knowledge and understanding*. While the verbs used here

bear some resemblance to those in Bloom's taxonomy of educational objectives (Bloom, 1956), an important distinction is that the processes in the financial literacy construct are not operationalised as a hierarchy of skills. They are, instead, parallel essential cognitive approaches, all of which are part of the financially literate individual's repertoire. The order in which the processes are presented here relates to a typical sequence of thought processes and actions, rather than to an order of difficulty or challenge. At the same time, it is recognised that financial thinking, decisions and actions are most often dependent on a recursive and interactive blend of the processes described in this section. For the purposes of the assessment, each task is identified with the process that is judged most central to its completion.

Identify financial information

This process is engaged when the individual searches and accesses sources of financial information, and identifies or recognises its relevance. In PISA 2012 the information is in the form of printed texts such as contracts, advertisements, charts, tables, forms and instructions. A typical task might ask students to identify the features of a purchase invoice, or recognise the balance on a bank statement. A more difficult task might involve searching through a contract that uses complex legal language to locate information that explains the consequences of defaulting on loan repayments. This process category is also reflected in tasks that involve recognising financial terminology, such as identifying "inflation" as the term used to describe increasing prices over time.

Example 5, *PAY SLIP*, shows an item that focuses on identifying and interpreting financial information.

Illustrative PISA Financial literacy item 5 – PAY SLIP

■ Figure 5.6 ■

Item for the unit *PAY SLIP*

Each month, Jane's employer pays money into Jane's bank account.

This is Jane's pay slip for July.

EMPLOYEE PAY SLIP: Jane Citizen		
Position: Manager	**1 July to 31 July**	
Gross salary	**2 800 zeds**	
Deductions	**300 zeds**	
Net salary	**2 500 zeds**	
Gross salary to date this year	**19 600 zeds**	

QUESTION

How much money did Jane's employer pay into Jane's bank account on 31 July?

A. 300 zeds

B. 2 500 zeds

C. 2 800 zeds

D. 19 600 zeds

Students are asked to *identify financial information* in a simple pay slip. The correct answer, 2 500 zeds, was selected by just over half of all students in the field trial.

Analyse information in a financial context

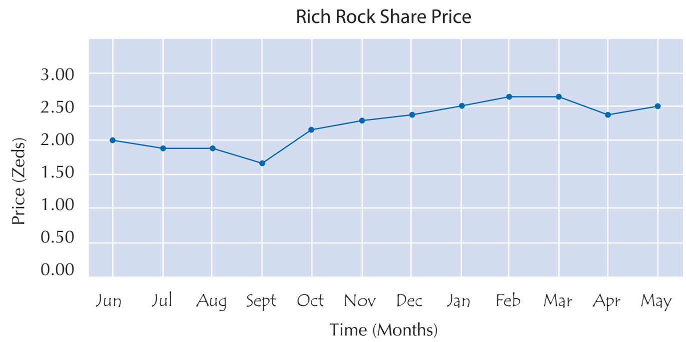
This process covers a wide range of cognitive activities undertaken in financial contexts, including interpreting, comparing and contrasting, synthesising, and extrapolating from information that is provided. Essentially it involves recognising something that is not explicit: identifying the underlying assumptions or implications of an issue in a financial context. For example, a task may involve comparing the terms offered by different mobile phone contracts, or working out whether an advertisement for a loan is likely to include unstated conditions. An example in this process category is provided below, in the unit *SHARES*.

Illustrative PISA Financial literacy item 6 – SHARES

■ Figure 5.7 ■

Item for the unit *SHARES*

This graph shows the price of one Rich Rock share over a 12-month period.



QUESTION

Which statements about the graph are true?

Circle "True" or "False" for each statement.

Statement	Is the statement true or false?
The best month to buy the shares was September.	True / False
The share price increased by about 50% over the year.	True / False

There are two parts to this question, which asks students to *analyse information in a financial context* by considering the information in a line graph about an investment product. The graph shows how the price of shares has changed over a year. The first part of the question assesses a student's understanding that shares should be bought when the price is low (in this case, September). The second part of the question assesses whether students can correctly identify the increase in share prices and calculate the percentage change over time. Just over half of all students correctly answered both parts of the question in the field trial, by circling "True" for the first statement and "False" for the second.

Evaluate financial issues

In this process the focus is on recognising or constructing financial justifications and explanations, drawing on financial knowledge and understanding applied in specified contexts. It involves such cognitive activities as explaining, assessing and generalising. Critical thinking is brought into play in this process, when students must draw on knowledge, logic and plausible reasoning to make sense of and form a view about a finance-related problem. The information that is required to deal with such a problem may be partly provided in the stimulus of the task, but students will need to connect such information with their own prior financial knowledge and understandings. In the PISA context, any information that is required to understand the problem is intended to be within the expected range of experiences of a 15-year-old – either direct experiences or those that can be readily imagined and understood. For example, it is assumed that 15-year-olds are likely to be able identify with the experience of wanting something that is not essential (such as a new sound system). A task based on this scenario could ask about the factors that might be considered in deciding on the relative financial merits of making a purchase or deferring it, given specified financial circumstances.

The next example task *AT THE MARKET* , based on the same stimulus as the illustrative PISA financial literacy item 1, asks students to evaluate information by drawing on everyday prior knowledge.

Illustrative PISA Financial literacy item 7 – AT THE MARKET

■ Figure 5.8 ■

Item for the unit *AT THE MARKET*

QUESTION

Buying a box of tomatoes may be a bad financial decision for some people.

Explain why ....

The purpose of this question is to recognise that buying in bulk may be wasteful for some people, or unaffordable in the short term. Students evaluate a financial issue in the situation presented and score full credit if they can explain that buying more tomatoes at a cheaper price may not always be a good financial decision. Answers such as the following, which refer to waste, received full credit:

- The tomatoes might rot before you use them all.
- Because you may not need 10 kg of tomatoes.

Another kind of answer that received full credit focused on individuals being unable to afford bulk buying:

- You have to spend 22 zeds (rather than 2.75 or 5.50 for 1 or 2 kg) and you might not have that amount to spend.
- You might have to go without something else that you need to pay for the box of tomatoes.

In the field trial, over 80% of students gained credit for this item, by referring to either waste or affordability.

NEW OFFER provides an example of a more demanding task that falls within the *evaluate financial issues* category.

Illustrative PISA Financial literacy item 8 – NEW OFFER

■ Figure 5.9 ■

Item for the unit *NEW OFFER*

Mrs Jones has a loan of 8 000 zeds with FirstZed Finance. The annual interest rate on the loan is 15%. Her repayments each month are 150 zeds.

After one year Mrs Jones still owes 7 400 zeds.

Another finance company called Zedbest will give Mrs Jones a loan of 10 000 zeds with an annual interest rate of 13%. Her repayments each month would also be 150 zeds.

QUESTION

What is one possible negative financial consequence for Mrs Jones if she agrees to the Zedbest loan?.....................................

....

NEW OFFER asks students to reflect on and evaluate the consequences of changing from one set of loan conditions to another – a context that is less likely to be familiar to 15-year-olds than the context provided in *AT THE MARKET*. In the case of *NEW OFFER*, all of the necessary information is provided in the question, but to gain credit students need to identify what is relevant and reflect on the consequences of taking a particular action. There are a number of kinds of responses that are awarded full credit. In the field trial, the most common credited response was that Mrs Jones would have more debt. Equally acceptable are responses that refer to specific conditions of the loan such as that the total interest paid (over the course of the loan) will be greater, that the length of the loan will be greater, and that there are possible fees associated with switching loan companies.

Just over 40% of students in the field trial gained credit for this item.

A third example of a task that fits within the *evaluate* process category is provided in the illustrative PISA financial literacy item 2, *SPENDING CHOICES* where students should draw on plausible reasoning in a financial context to assess which of the tasks listed require prompt attention.

Apply financial knowledge and understanding

The fourth process picks up a term from the definition of financial literacy: "to apply such [financial] knowledge and understanding". It focuses on taking effective action in a financial setting by using knowledge of financial products and contexts, and understanding of financial concepts. This process is reflected in tasks that involve performing calculations and solving problems, often taking into account multiple conditions. An example of this kind of task is calculating the interest on a loan over two years. This process is also reflected in tasks that require recognition of the relevance of prior knowledge in a specific context. For example, a task might require the student to work out whether purchasing power will decline or increase over time when prices are changing at a given rate. In this case, knowledge about inflation needs to be applied.

The following example, *TRAVEL MONEY*, falls into the process category *apply financial knowledge and understanding*.

Illustrative PISA Financial literacy item 9 – TRAVEL MONEY

■ Figure 5.10 ■

Item for the unit *TRAVEL MONEY*

Natasha works in a restaurant 3 evenings each week.

She works for 4 hours each evening and she earns 10 zeds per hour.

Natasha also earns 80 zeds each week in tips.

Natasha saves exactly half of the total amount of money she earns each week.

QUESTION

Natasha wants to save 600 zeds for a holiday.

How many weeks will it take Natasha to save 600 zeds?..

This task requires students to consider a set of conditions and constraints, while planning ahead to pay for future expenses – working out how long it will take to save for a holiday, given a fixed amount of savings each week. The correct answer is "6 weeks". Fewer than half of the field trial sample gained credit for this item.

Contexts

In building a framework, and developing and selecting assessment items based on this framework, attention is given to the breadth of contexts in which the domain literacy is exercised. Decisions about financial issues are often dependent on the contexts or situations in which they are presented. By situating tasks in a variety of contexts the assessment offers the possibility of connecting with the broadest possible range of individual interests across a variety of situations in which individuals need to function in the 21st century.

Certain situations will be more familiar to 15-year-olds than others. In PISA, assessment tasks are framed in situations of general life, which may include but are not confined to school contexts. The focus may be on the individual, family or peer group, on the wider community, or even more widely on a global scale.

As a starting point, the FEG looked at the contexts used in the Programme for the International Assessment of Adult Competencies (PIAAC) literacy framework: *education and work, home and family, leisure and recreation,* and *community and citizenship* (OECD, 2009b). For the purposes of the financial literacy domain, the heading *leisure and recreation* was replaced by *individual* to reflect the fact that many of the financial interactions that young people have are related to themselves as individual consumers, using products such as mobile phones or laptops, as well as accessing leisure facilities or funding recreation. It was further decided to replace *community and citizenship* with *societal*. While *community and citizenship* captures the idea of a perspective wider than the personal, it was felt that the term *community* was not wide enough. *Societal*, by contrast, implicitly encompasses national and global situations as well as the more local, thus better fitting the potential reach of financial literacy. The contexts identified for the PISA financial literacy assessment are, then, *education and work, home and family, individual* and *societal*.

Education and work

The context of *education and work* is of great importance to young people. The educational context is obviously relevant to PISA students, since they are by definition a sample of the school-based population; indeed, many of them will continue in education or training for some time. However, many other 15-year-olds move from school into the labour force within one to two years, and many 15-year-old students are engaged in casual employment outside school hours. Therefore, both currently and for the medium term, the occupational context is also relevant for PISA students. Virtually all 15-year-olds will be starting to think about financial matters related to both education and work, whether they are spending existing earnings, considering future education options or planning their working life.

Typical tasks within this context could include understanding payslips, planning to save for tertiary study, investigating the benefits and risks of taking out a student loan, and participating in workplace savings schemes.

Item 5, in *PAY SLIP*, and item 9, in *TRAVEL MONEY*, illustrate the kind of task designed to reflect the *education and work* context category. Specifically, they are examples the *work* context, asking students to address financial problems related to earned income; the first asking them to identify information on a payslip, and the second using the context of income to make a savings plan.

Home and family

Home and family includes financial issues relating to the costs involved in running a household. Family is the most likely household circumstance for 15-year-olds; however, this category also encompasses households that are not based on family relationships, such as the kind of shared accommodation that young people often use shortly after leaving the family home. Tasks within this context may include buying household items or family groceries, keeping records of family spending and making plans for family events. Decisions about budgeting and prioritising spending may also be framed within this context.

The two items from the *AT THE MARKET* unit, items 1 and 7 are categorised as home and family, since grocery shopping is usually done for a household. Item 2, in *SPENDING CHOICES*, is also in this context category: the setting is shared accommodation, and the choices to be made will affect the household (in this case, friends rather than a family sharing accommodation).

Individual

The context of the *individual* is important within personal finance since there are many decisions that a person takes entirely for personal benefit or gratification, and many risks and responsibilities that must be borne by individuals. Decisions taken that fit within this context include choosing personal products and services such as clothing, toiletries or haircuts, or buying consumer goods such as electronic or sports equipment, as well as commitments such as season tickets or a gym membership. These decisions span essential personal needs, as well as leisure and recreation. Although the decisions made by an individual may be influenced by the family and society, when it comes to opening a bank account or getting a loan it is the individual who has the legal responsibility for such decisions. The context *individual* therefore includes contractual issues around events such as opening a bank account, purchasing consumer goods, paying for recreational activities, and dealing with relevant financial services that are often associated with larger consumption items, such as credit and insurance.

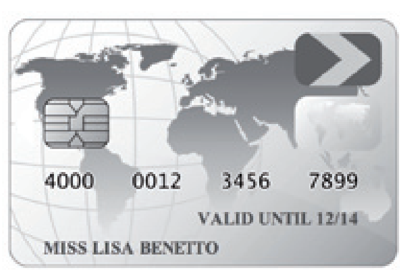
NEW BANK CARD is an example of an item from the *individual* context category.

Illustrative PISA Financial literacy item 10 – NEW BANK CARD

■ Figure 5.11 ■

Item for the unit *NEW BANK CARD*

Lisa lives in Zedland. She receives this new bank card.



QUESTION

The following day, Lisa receives the Personal Identification Number (PIN) for the bank card.

What should Lisa do with the PIN?

A. Write the PIN on notepaper and keep this in her wallet.
B. Tell the PIN to her friends.
C. Write the PIN on the back of the card.
D. Memorise the PIN.

This task assesses students' understanding of the individual's responsibility in maintaining security when accessing and using electronic banking. This question asks students to evaluate which of the four presented options is best practice when using a bank card. Over 90% of students in the field trial chose the correct option of memorising the PIN (D).

Other items from the *individual* context category, shown earlier, include item 3, in *MOTORBIKE INSURANCE*, item 6, in *SHARES*, and item 8, in *NEW OFFER,* all of which illustrate decisions that impact on the individual (renewing insurance, buying shares, refinancing a loan).

Societal

The environment young people are living in is characterised by change, complexity and interdependence. Globalisation is creating new forms of interdependence where actions are subject to economic influences and consequences that stretch well beyond the individual and the local community. While the core of the financial literacy domain is focused on personal finances, the *societal* context recognises that individual financial well-being cannot be entirely separated from the rest of society. Personal financial well-being affects and is affected by the local community, the nation and even global activities. Financial literacy within this context includes matters such as being informed about consumer rights and responsibilities, understanding the purpose of taxes and local government charges, being aware of business interests, and taking into account the role of consumer purchasing power. It extends also to considering financial choices such as donating to non-profit organisations and charities.

The task *BANK ERROR* (item 4, shown earlier) is categorised as falling within the *societal* context, since it relates to fraudulent behaviour targeted across society.

Non-cognitive factors

The PISA working definition of financial literacy includes the non-cognitive terms *motivation* and *confidence*, attitudes which, according to some, have an influence on money management behaviour (Johnson and Staten, 2010). PISA conceives of both financial attitudes and behaviour as aspects of financial literacy in their own right. Attitudes and behaviour are also of interest in terms of their interactions with the cognitive elements of financial literacy. Information collected about the financial attitudes and behaviour of 15-year-olds will also potentially constitute useful baseline data for any longitudinal investigation of the financial literacy of adults, including their financial behaviours.

The FEG identified four non-cognitive factors for inclusion in the framework: *access to information and education, access to money and financial products, attitudes towards and confidence about financial matters,* and *spending and saving behaviour.*[2]

Access to information and education

There are various sources of financial information available to students, including friends, parents or other family members. It is useful to know which sources of information are accessed most frequently and to ascertain whether higher levels of financial literacy are associated with particular sources of information. Policy makers can also use this information to ascertain how well messages about financial issues are being communicated, and where to target new interventions.

The education and training received by students also varies within and across countries. Information about the extent to which there is a link between levels of financial literacy and financial education inside and outside schools is likely to be particularly useful in shaping education programmes for improving financial literacy.

Access to money and financial products

Students who have had more personal experience dealing with financial matters might be expected to perform better on the cognitive assessment. Those who regularly make decisions about how to manage their own money are likely to know more about financial matters, even if they have not had specific instruction, than those who do not. That experience may come from earning money, from using financial products such as credit and debit cards, or from dealing with the banking system. A key policy question in this area is, "To what extent do real-life experiences of the financial world influence young people's financial literacy?"

Attitudes towards and confidence about financial matters

Attitudes are considered important constituents of financial literacy. Moreover, individual preferences are important determinants of financial behaviour and can interact with financial literacy. It is hypothesised that research from behavioural psychology may yield interesting results with regard to financial literacy, and better inform policy makers trying to improve the efficiency of programmes. Areas identified for possible investigation include risk tolerance – a willingness to accept the possibility of a loss in order to achieve greater gain (Barsky et al., 1997; Holt and Laury, 2002); and time sensitivity – willingness to trade immediate reward for greater gain at a future date (Barsky et al., 1997; Holt and Laury, 2002).

Spending and saving behaviour

While items on the cognitive assessment test students' ability to make particular spending and savings decisions, it is also useful to have some measure of what their actual (reported) behaviour is: that is, how students save and spend in practice. PISA financial literacy will provide evidence on the relationship between financial literacy knowledge and financial behaviour, by looking at the relationship between 15-year-olds' reported behaviour and their results on the cognitive financial literacy assessment.

ASSESSING FINANCIAL LITERACY

The previous section has outlined the conceptual framework for financial literacy. The concepts in the framework must in turn be represented in tasks and questions in order to collect evidence of students' proficiency in financial literacy. In this section we discuss the structure of the assessment, the distribution of tasks across the framework variables, and the choice of response formats. This is followed by a short discussion of the impact of knowledge and skills from other domains on financial literacy and the implications for the assessment. To conclude, we describe the method by which data about financial behaviours and experience will be collected.

The conceptual framework is concerned with mapping the domain, not just for the 2012 assessment, but more broadly. It lays out the definition and the major variables that are addressed in the assessment instrument. The key ideas have been elaborated through lists of sub-topics and examples in the preceding section. These elaborations should not be interpreted as a checklist of tasks included in the 2012 assessment. Given that only one hour of financial literacy assessment material is being administered in PISA 2012, there is not enough space to cover every detail of each variable. It is anticipated that further aspects of the domain will be included in assessment tasks in future administrations.

The structure of the assessment

The PISA paper-and-pen assessment is designed as a two-hour test comprising four 30-minute clusters of test material from one or more cognitive domains. Financial literacy was allocated two clusters (that is, 60 minutes of testing time) in the 2012 main survey. Analysis of completion rates in the field trial was used to determine that the vast majority of students could be expected to complete 20 financial literacy items within 30 minutes. Accordingly, from the 75 financial literacy tasks administered in the field trial, 40 were selected for the main survey.

In the 2012 main survey each test booklet that includes the two clusters of financial literacy items also includes one cluster of mathematics test items and one cluster of reading items. To reduce any effects from the order of the clusters within a booklet, four test booklets containing financial literacy clusters have been created, with the financial literacy, mathematics and reading clusters appearing in different positions.

As with other PISA assessment domains, financial literacy items are grouped in units comprising one or two items based around a common stimulus. The selection includes financially-focused stimulus material in diverse formats, including prose, diagrams, tables, charts and illustrations.

The assessment comprises a broad sample of items covering a range of difficulty that will enable the strengths and weaknesses of students and key subgroups to be measured and described.

Response formats and coding

Some PISA paper-and-pen items require short handwritten responses, others responses of one or two sentences, whilst others can be answered by circling an answer or ticking a box. Decisions about the form in which the data are collected – the response formats of the items – are based on what is considered appropriate given the kind of evidence that is being collected, and also on technical and pragmatic considerations. In the financial literacy assessment as in other PISA assessments, two broad types of items are used: constructed-response items and selected-response items.

Constructed-response items require students to generate their own answers. The format of the answer may be a single word or figure, or may be longer: a few sentences or a worked calculation. Constructed-response items that require a more extended answer are ideal for collecting information about students' capacity to explain decisions or demonstrate a process of analysis. Item 9, in *TRAVEL MONEY*, illustrates a constructed response item that calls for a single figure, where there is a very restricted range of credit-worthy responses. Items 1 and 7, both from the unit *AT THE MARKET*, and item 8, in *NEW OFFER*, are typical of tasks that require more extended responses, where many different kinds of answers may gain full credit.

In the scoring of the *NEW OFFER* task, for example, four different kinds of full credit responses are identified:

(*i*) Answers that refer to Mrs Jones having more debt if she takes on the new loan offer:

- She will owe more money.
- She will be unable to control her spending.
- She is going deeper into debt.

(*ii*) Answers that refer to Mrs Jones having to pay more interest:

- 13% of 10 000 is greater than 15% of 8 000.

(*iii*) Answers that refer to the increased length of time over which Mrs Jones will have the debt:

- It might take longer to repay because the loan is bigger and the payments are the same.

(*iv*) Answers that refer to the possibility that Mrs Jones will have to pay a cancellation fee if she cancels her initial loan agreement with FirstZed Finance:

- She may have a penalty fee for paying the FirstZed loan early.

The second broad type of item, with regard to format and coding, is selected response. This kind of item requires students to choose one or more alternatives from a given set of options. The most common type in this category is the simple multiple-choice item, which requires the selection of one from a set of (usually) four options: see item 5, in *PAY SLIP*, and item 10, in *NEW BANK CARD*. A second type of selected-response item is complex multiple choice, in which students respond to a series of "Yes/No"-type questions. Item 2, in *SPENDING CHOICES*, illustrates a "Yes/No" set of selections. For this task students need to make three independent correct selections to gain credit. Item 3, in *MOTORBIKE INSURANCE*, has a similar format, in that three independent correct selections must be made to gain credit; but in this case each selection is from three options: "Increases cost", "Reduces cost" and "Has no effect on cost". Selected-response items are typically regarded as most suitable for assessing items associated with identifying and recognising information, but they are also a useful way of measuring students' understanding of higher-order concepts that they themselves may not easily be able to express.

Although particular item formats lend themselves to specific types of questions, care needs to be taken that the format of the item does not affect the interpretation of the results. Research suggests that different groups (for example, boys and girls, and students in different countries) respond differentially to the various item formats. Several research studies on response format effects based on PISA data suggest that there are strong arguments for retaining a mixture of multiple-choice and constructed-response items. In their study of PISA reading literacy compared with the IEA Reading Literacy Study (IEARLS), Lafontaine and Monseur (2006) found that response format had a significant impact on gender performance. In another study, countries were found to show differential equivalence of item difficulties in PISA reading on items in different formats (Grisay and Monseur, 2007). This finding may relate to the fact that students in different countries are more or less familiar with the particular formats. In summary, the PISA financial literacy option includes items in a variety of formats to minimise the possibility that the item format influences student performance. Such an influence would be extrinsic to the intended object of measurement: in this case, financial literacy.

When considering the distribution of item formats, the question of resources must be weighed as well as the equity issues discussed in the preceding paragraphs. All except the most simple of constructed-response items are coded by expert

judges who must be trained and monitored. Selected response and very short "closed" constructed response items do not require expert coding and therefore demand fewer resources.

The proportions of constructed- and selected-response items are determined taking account of all these considerations. The majority of the items selected for the PISA 2012 main survey do not require expert judgement.

Most items are coded dichotomously (full credit or no credit), but where appropriate an item's coding scheme allows for partial credit. Partial credit makes possible more nuanced scoring of items. Some answers, even though incomplete, are better than others. If incomplete answers for a particular question indicate a higher level of financial literacy than inaccurate or incorrect answers, a scoring scheme has been devised that allows partial credit for that question. Such "partial credit" items yield more than one score point.

Distribution of score points

In this section we outline the distribution of score points across the categories of the three main framework characteristics discussed previously. The term "score points" is used in preference to "items", as some partial credit items are included. The distributions are expressed in terms of ranges, indicating the approximate weighting of the various categories.

While each PISA financial literacy item is categorised according to a single content, a single process and a single context category it is recognised that, since PISA aims to reflect real-life situations and problems, often elements of more than one category are present in a task. In such cases, the item is identified with the category judged most integral to responding successfully to the task.

The target distribution of score points according to financial literacy content areas is shown in Table 5.1.

Table 5.1
Approximate distribution of score points in financial literacy, by content

Money and transactions	Planning and managing finances	Risk and reward	Financial landscape	Total
30% - 40%	25% - 35%	15% - 25%	10% - 20%	100%

The distribution reflects that *money and transactions* is considered to be to the most immediately relevant content area for 15-year-olds.

Table 5.2 shows the target distribution of score points by the four processes.

Table 5.2
Approximate distribution of score points in financial literacy, by processes

Identify financial information	Analyse information in a financial context	Evaluate financial issues	Apply financial knowledge and understanding	Total
15% - 25%	15% - 25%	25% - 35%	25% - 35%	100%

The weighting shows that greater importance was attributed to evaluating financial issues and applying financial knowledge and understanding.

Table 5.3 shows the target distribution of score points by the four contexts.

Table 5.3
Approximate distribution of score points in financial literacy, by contexts

Education and work	Home and family	Individual	Societal	Total
10% - 20%	30% - 40%	35% - 45%	5% -15%	100%

Consistent with an assessment of personal financial literacy of 15-year-olds, there is a clear emphasis on *individual*, but also a weighting towards the financial interests of the household or family unit. *Education and work* and *societal* contexts are given less emphasis, but included in the scheme as they are important elements of financial experience.

The impact of other domain knowledge and skills on financial literacy

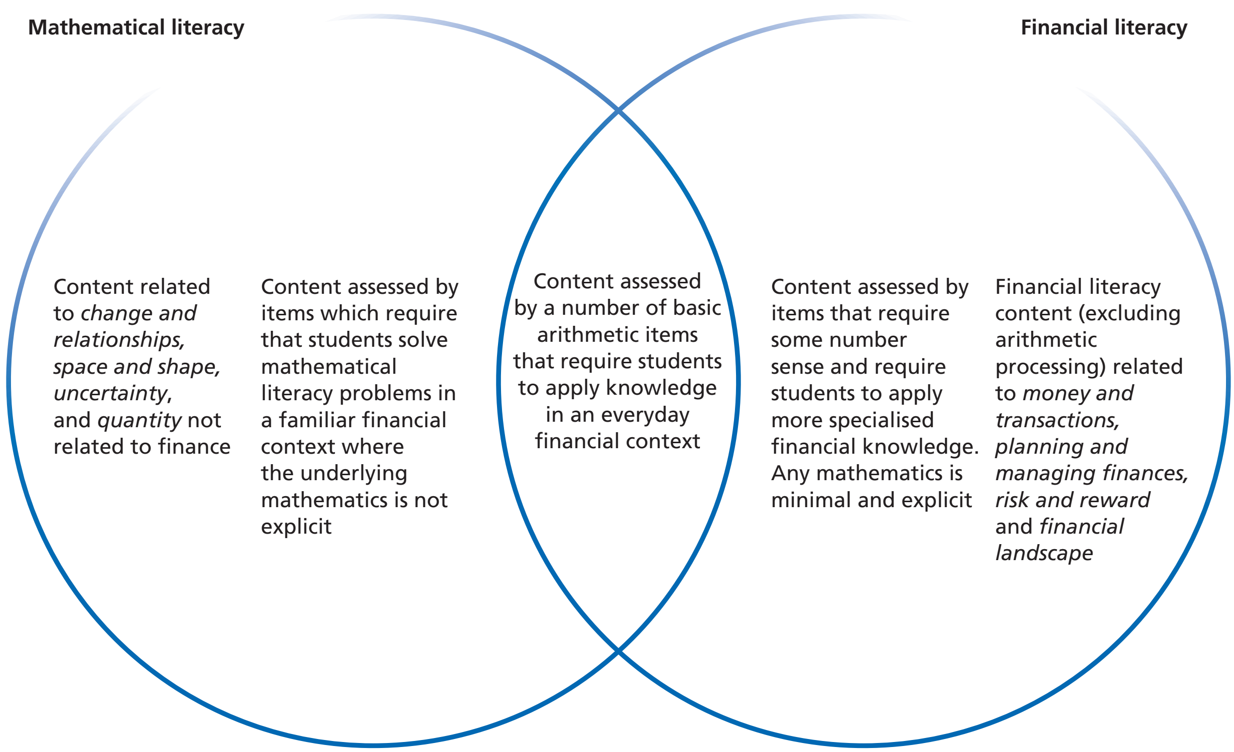
Numeracy skills

A certain level of numeracy (or mathematical literacy) is regarded as a necessary condition of financial literacy. Huston (2010) argues that "if an individual struggles with arithmetic skills, this will certainly impact his/her financial literacy. However, available tools (e.g. calculators) can compensate for these deficiencies; thus, information directly related to successfully navigating personal finances is a more appropriate focus than numeracy skills for a financial literacy measure". It is therefore common for financial literacy assessments to include items with a mathematical literacy aspect, even though that aspect is not the primary focus of the whole measure. Lusardi et al. (2010) reported that, in the 1997 National Longitudinal Survey of Youth conducted in the United States, three financial literacy questions "differentiated well between naive and sophisticated respondents." Two of the three questions, on interest rates and inflation, required some basic competence in mathematical literacy. Mathematically-related proficiencies such as number sense, familiarity with multiple representations of numbers, and skills in mental calculation, estimation, and the assessment of reasonableness of results are intrinsic to some aspects of financial literacy.

On the other hand there are large areas where the content of mathematical literacy and financial literacy do not intersect. As defined in the PISA 2012 mathematics literacy framework, mathematical literacy defines four content areas: *change and relationships, space and shape, quantity* and *uncertainty*. Of these, only *quantity* directly intersects with the content of the PISA financial literacy assessment. Unlike the mathematical literacy content area *uncertainty*, which requires students to apply probability measures and statistics, in the PISA assessment the financial literacy content area *risk and reward* requires an understanding of the features of a particular situation or product that indicate risk/reward. This is a non-numeric appreciation of the way financial well-being can be affected by chance and an awareness of the related products and actions to protect against loss. In the financial literacy assessment, the quantity-related proficiencies listed previously can be applied to problems requiring more financial knowledge than can be expected in the mathematical literacy assessment. Similarly, knowledge about financial matters and capability in applying such knowledge and reasoning in financial contexts (in the absence of any specifically mathematical content) characterise much of all four content areas of financial literacy: *money and transactions, planning and managing finances, risk and reward* and *financial landscape*. Figure 5.12 represents the relationship between the content of mathematical literacy and financial literacy in PISA.

■ Figure 5.12 ■

Relationship between the content of financial literacy and mathematical literacy in PISA



Operationally, there are few items populating the portion of the diagram where the two circles intersect. In the financial literacy assessment, the nature of the mathematical literacy expected is basic arithmetic: the four operations (addition, subtraction, multiplication and division) with whole numbers, decimals and common percentages. Such arithmetic occurs as an intrinsic part of the financial literacy context and enables financial literacy knowledge to be applied and demonstrated. Item 1, from the unit *AT THE MARKET*, illustrates an item that requires such arithmetic skills: the mathematics involved (dividing by a factor of ten) is very basic and is of a level well within the reach of most 15-year-olds. Use of financial formulae (requiring capability with algebra) is not considered appropriate. Dependence on calculation is minimised in the assessment; tasks are framed in such as way as to avoid the need for substantial or repetitive calculation. The calculators used by students in their classrooms and on the PISA mathematics assessment will also be available in the financial literacy assessment, but success in the items will not depend on calculator use.

Reading and vocabulary

It is assumed that all students taking part in the financial literacy assessment will have some basic reading proficiency, even while it is known from previous PISA surveys that reading skill varies widely both within and across countries (OECD, 2010b). To minimise the level of reading literacy required, stimulus material and task statements are generally designed to be as clear, simple and brief as possible. In some cases, however, stimulus may deliberately present complex or somewhat technical language: the capacity to read and interpret the language of financial documents or pseudo financial documents is regarded as part of financial literacy. In item 4, in *BANK ERROR*, attentive reading of the e-mail message is required to detect the likelihood that the message is part of a scam.

Highly technical terminology relating to financial matters is avoided. The FEG has advised on terms that it judges reasonable to expect 15-year-olds to understand. Some of these terms may be the focus of assessment tasks. For example, *PAY SLIP* (item 5), assesses whether students are able to read a simple pay slip and recognise (or infer the meaning of) the terms “gross” and “net”.

Collecting data about financial behaviour and experience

Information about non-cognitive factors related to financial literacy is collected in a short student questionnaire at the end of the cognitive assessment of financial literacy. Items address aspects of three of the four key areas identified for inclusion by the financial literacy expert group: *access to information and education; access to money and financial products;* and *spending and saving behaviour*. The questionnaire comprises a small set of questions that explore the range and types of students’ interest in and experience with financial matters.

The questions for the short questionnaire are based on questions from existing national surveys of financial literacy. Additional information that is pertinent to understanding the distribution of financial literacy can be drawn from the standard PISA student background questionnaires. Data about the students’ home situation (family socioeconomic status in particular) and school experience may be relevant to understanding their financial literacy results. In addition, the school questionnaire, which heads of all schools in the PISA sample are asked to complete, includes questions about the availability of financial education for their students, and access to professional development in financial education for their teachers.

REPORTING FINANCIAL LITERACY

The data from the financial literacy assessment is held in a database separate from the main PISA database. This database includes, for the sampled students, their financial literacy, mathematics and reading cognitive results, the behaviour data from the short questionnaire on financial literacy, and data from the general student questionnaire and school questionnaire.

It is therefore possible to report on financial literacy as an independent result, and on financial literacy in relation to mathematics performance, reading performance, financial behaviour, and in relation to some background variables such as socioeconomic status and immigrant status. The results also allow the development of further work under the aegis of the OECD Project on Financial Education.

The financial literacy cognitive data is scaled in a similar way to the other PISA data. A comprehensive description of the modelling technique used for scaling can be found in the *PISA 2006 Technical Report* (OECD, 2009c).

Each item is associated with a particular point on the PISA financial literacy scale that indicates its difficulty, and each student's performance is associated with a particular point on the same scale that indicates the student's estimated proficiency.

As with the other PISA domains, the relative difficulty of tasks in a test is estimated by considering the proportion of test takers getting each question correct. The relative proficiency of students taking a particular test is estimated by considering the proportion of test items that they answer correctly. A single continuous scale showing the relationship between the difficulty of items and the proficiency of students will be constructed.

The scale is divided into levels, according to a set of statistical principles, and then descriptions are generated based on the tasks that are located within each level, to encapsulate the kinds of skills and knowledge needed to successfully complete those tasks. The scale and set of descriptions are known as a described proficiency scale.

By calibrating the difficulty of each item, it is possible to locate the degree of financial literacy that the item represents. By showing the proficiency of each student on the same scale, it is possible to describe the degree of financial literacy that the student possesses. The described proficiency scale helps in interpreting what students' financial literacy scores mean in substantive terms.

Following PISA practice, a scale is being constructed having a mean of 500 and a standard deviation of 100 (based on the participation of 13 OECD countries). Given the number of items in the 2012 assessment (40), four levels of proficiency in financial literacy will be described, as a first step in reporting how competency in financial literacy develops, and to enable comparisons of student performance between and within participating countries and economies. The optional assessment of financial literacy in PISA 2012 will provide essential inputs and data for both PISA and the OECD Project on Financial Education.

Notes

1. The Joint Ministerial Statement from the 2012 APEC Finance Ministerial Meeting is available at *http://www.apec.org/Meeting-Papers/Ministerial-Statements/Finance/2012_finance.aspx.*

2. *Attitudes towards* and *confidence about financial matters* are not covered in the PISA 2012 financial literacy assessment.

References

Atkinson, A. and **F. Messy** (2012), "Measuring Financial Literacy: Results of the OECD / International Network on Financial Education (INFE) Pilot Study", *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 15, OECD Publishing.

Barsky, R.B., F.T. Juster, M.S. Kimball and **M.D. Shapiro** (1997), "Preference parameters and behavioural heterogeneity: An experimental approach in the health and retirement study", *Quarterly Journal of Economics*, No. 11, pp. 537-539.

Bernheim, D., D. Garrett and **D. Maki** (2001), "Education and saving: The long-term effects of high school financial curriculum mandates", *Journal of Public Economics*, No. 85, pp. 435-565.

Bloom, B.S. (ed.) (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, David McKay, New York.

Cole, S., T. Sampson and **B. Zia** (2011), "Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets?", *The Journal of Finance*, Vol. 66, No. 6, pp. 1933-1967.

G20 (2012), *G20 Leaders Declaration, Los Cabos, www.g20mexico.org/images/stories/docs/g20/conclu/G20_Leaders_Declaration_2012.pdf.*

Gerardi, K., L. Goette, and **S. Meier** (2010), "Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data", Federal Reserve Bank of Atlanta, Working Paper Series No. 2010-10.

Grisay, A. and **C. Monseur** (2007), "Measuring the equivalence of item difficulty in the various versions of an international test", *Studies in Educational Evaluation*, Vol. 33, No. 1, pp. 69-86.

Habschick, M., B. Seidl and **J. Evers** (2007), *Survey of Financial Literacy Schemes in the EU27*, Hamburg.

Hastings, J. and **L. Tejeda-Ashton** (2008), Financial Literacy, Information, and Demand Elasticity: Survey and Experimental Evidence from Mexico, NBER Working Paper, 14538, Cambridge, Massachusetts.

Hilgert, M.A., J.M. Hogarth and **S.G. Beverly** (2003), Household Financial Management: The Connection Between Knowledge and Behavior, *Federal Reserve Bulletin*, Vol. 89, No. 7, pp. 309-322.

Holt, C. and **S. Laury** (2002), "Risk aversion and incentive effects", *American Economic Review*, Vol. 92, No. 5, pp.1644-1655.

Huston, S.J. (2010), "Measuring financial literacy", *The Journal of Consumer Affairs*, Vol. 44, No. 2, pp. 296-316.

Johnson, C. and **M. Staten** (2010), *Do inter-temporal preferences trump financial education courses in driving borrowing and payment behaviour?,* Paper presented at the 1st Annual Boulder Conference on Consumer Financial Decision Making, June 27-29, 2010, Boulder, Colorado.

Kempson, E., S. Collard and **N. Moore** (2005), *Measuring financial capability: an exploratory study*, Financial Services Authority, London.

Lafontaine, D. and **C. Monseur** (2006), *Impact of Test Characteristics on Gender Equity Indicators in the Assessment of Reading Comprehension*, University of Liège.

Lusardi, A. (2009), U.S. Household Savings Behavior: The Role of Financial Literacy, Information and Financial Education Programs, in C. Foote, L. Goette and S. Meier (eds.), *Policymaking Insights from Behavioral Economics*, Federal Reserve Bank of Boston, pp. 109-149.

Lusardi, A. and **O.S. Mitchell** (2008), "Planning and Financial Literacy: How Do Women Fare? ", *American Economic Review*, Vol. 98, No. 2, pp. 413-417.

Lusardi, A. and **O.S. Mitchell** (2011), Financial Literacy and Planning: Implications for Retirement Wellbeing, in A. Lusardi and O. S. Mitchell (eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, Oxford University Press.

Lusardi, A., O.S. Mitchell and **V. Curto** (2010), "Financial literacy among the young", *The Journal of Consumer Affairs,* Vol. 44, No. 2, pp. 358-380.

Lusardi, A. and **P. Tufano** (2009a), "Debt Literacy, Financial Experiences, and Overindebtedness", NBER Working Paper No. 14808, Cambridge, Massachusetts.

Lusardi, A. and **P. Tufano** (2009b), Teach Workers about the Perils of Debt, *Harvard Business Review* (November), pp. 22-24.

Ministerial Council for Education Early Childhood Development and Youth Affairs (2011), National Consumer and Financial Literacy Framework, MCEETYA Consumer and Financial Literacy Working Party, Melbourne.

Moore, D. (2003), *Survey of Financial Literacy in Washington State: Knowledge, Behavior, Attitudes, and Experiences*, Social and Economic Sciences Research Center, Washington State University.

OECD (2005a), *Recommendation on Principles and Good Practices for Financial Education*, OECD Publishing.

OECD (2005b), *Improving Financial Literacy: Analysis of Issues and Policies*, OECD Publishing.

OECD (2008), *Improving Financial Education and Awareness on Insurance and Private Pensions*, OECD Publishing.

OECD (2009a), *Financial Literacy and Consumer Protection: Overlooked Aspects of the Crisis*, OECD Publishing.

OECD (2009b), "PIAAC Literacy: A Conceptual Framework", *OECD Education Working Paper* No. 34, OECD Publishing.

OECD (2009c), *PISA 2006 Technical Report*, OECD Publishing.

OECD (2010a), *PISA 2009 Framework: Key Competencies in Reading, Mathematics and Science*, OECD Publishing.

OECD (2010b), *PISA 2009 Results: What Students Know and Can Do,* Vol. I, OECD Publishing.

OECD (2011), *Education at a Glance 2011: OECD Indicators*, OECD Publishing.

OECD (forthcoming), *Financial Education in Schools: Challenges, Case Studies and Policy Guidance*, OECD Publishing.

OECD INFE (2009), *Financial Education and the Crisis: Policy Paper and Guidance*, OECD Publishing.

OECD INFE (2011), *Measuring Financial Literacy: Core Questionnaire in Measuring Financial Literacy: Questionnaire and Guidance Notes for conducting an Internationally Comparable Survey of Financial literacy*, OECD Publishing.

OECD INFE (2012), *OECD/INFE High-Level Principles on National Strategies for Financial Education*, OECD Publishing.

Schuchardt, J., S.D. Hanna, T.K. Hira, A.C. Lyons, L. Palmer and **J.J. Xiao** (2009), "Financial Literacy and Education Research Priorities", *Journal of Financial Counseling and Planning*, Vol. 20, No. 1, pp. 84-95.

Smithers, R. (2010), University students expect to graduate with debts in excess of £15,000, *The Guardian*, 18 March 2010, *www.guardian.co.uk/money/2010/mar/18/university-students-graduate-mouting-debts*.

Stango, V. and **J. Zinman** (2009), "Exponential Growth Bias and Household Finance", *Journal of Finance*, Vol. 64, No. 6, pp. 2807-2849.

Van Rooij, M.A., A. Lusardi and **R. Alessie** (2011), "Financial Literacy and Stock Market Participation", *Journal of Financial Economics*, Vol. 101, No. 2, pp.449-472.

Yoong, J. (2011), Financial Illiteracy and Stock Market Participation: Evidence from the RAND American Life Panel, in A. Lusardi and O. S. Mitchell (eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, Oxford University Press.

6

PISA 2012 Context Questionnaires Framework

The PISA 2012 context questionnaire framework describes how PISA can be developed further as a sustainable database for educational policy, and research. The framework maps out a design for the PISA context questionnaires that will be sustainable into the future. It also makes recommendation regarding aspects of design and analysis that are intended to build on the existing strengths of PISA and mapping an even stronger path into the future. Finally, the framework provides the theoretical and scientifically rigorous underpinnings of the concepts, scales and indices developed for the PISA 2012 cycle and that will be constructed using the different context questionnaires implemented in participating countries and economies. In particular, the framework discusses in detail new individual outcome measures – strategies, beliefs and motivation – related to mathematical literacy and innovative indicators of opportunity to learn and quality of instruction at the level of individual students, schools and systems.

INTRODUCTION

Since its inception in the late 1990s, the OECD Programme for International Student Assessment (PISA) has been known for its important contribution to education policy discussions within OECD and partner countries and economies. Literacy and life skills of adolescents, assessed at the age of 15, capture the yield of years of learning both within and outside schools. Reading, science and mathematical literacy are important prerequisites for individual success in academic, professional, civic, social and private contexts. When reported by systems or countries, the average level of those skills is an important indicator of human capital, which in turn has an impact on the prosperity and well-being of society as a whole. How these skills are distributed throughout the population, dependent on gender, origin, socio-economic background or geographic region, will help to understand diversity and equity within and between countries. The performance data delivered by PISA have been combined with information on student backgrounds to provide information on level, diversity and equity of life skills. Policy makers, researchers, educators, parents, students and the general public alike can compare what their respective country has achieved to other countries, using PISA to set benchmarks. In order to elicit the necessary background information, students participating in any of the PISA cycles have been questioned for example about their parents' education and occupation as well as their own immigration background and gender.

Literacy and life skills are to a large extent products of family and school education. In order to describe these learning contexts and to help understand their impact on student performance, PISA has also asked students, school principals, and sometimes parents about their perspectives on the learning context. This additional information allows PISA to answer questions such as: Do countries differ in terms of instructional culture, i.e. can we identify country-specific profiles of teaching practices? Do different schools within countries serve different student populations and, if so, are their policies and practices targeted towards their respective populations? Can performance differences between individual students be accounted for by differences in societal characteristics, families, school resources, instructional practices and communities? The answers suggested by PISA have been highly influential in debates on quality, efficacy and efficiency of education systems around the globe. Within one decade, PISA 2000, 2003, 2006 and 2009 have established one of the most important global databases that may be used to draw policy insights and assess educational effectiveness. Policy makers may use the answers they obtain from PISA to guide their decision making by re-defining goals, setting priorities for educational reform, re-allocating resources, promoting new pedagogical concepts and organisational strategies, and ultimately improving education. As it re-evaluates education systems every three years, PISA both informs and monitors change in education.

Given that PISA is entering its fifth cycle of surveys, it is time to reconsider the overarching framework that guides the development of questionnaires and the way that this information is used in analysing and reporting data. Between 2000 and 2009, the PISA student questionnaires, for instance, have gathered data on more than 70 indicators (scales aggregated from the questionnaire data), covering family resources and activities, the learning environment at school and in the classroom, and students' beliefs and motivations. However, none of these scales has been administered in all four cycles, while the majority (more than 55 indicators) have been used just once. Thus, each of the PISA cycles has been to some extent designed as an independent study on learning contexts that focuses on a specific major domain. Questionnaire content has largely been determined by contemporary priorities rather than long-term policy and research goals. Undoubtedly, this design principle was useful as long as the survey was still being developed and new domains or innovative measures continued to emerge.

Certainly, the design and analytical framework of the survey will continue to adapt to changes in society and education, such as the increasing importance of information and communication technology (ICT) both as a learning tool and a domain of cross-curricular skills, or the growing need to combine multiple settings for student learning (schools, after-school programmes, e-learning, or even home schooling). At the same time, however, PISA has matured. As a sound and stable basis for international comparative studies and trend information on education systems, PISA requires a coherent architecture for future context questionnaires that balances stability with innovation, and domain-specific with general issues. This document aims to develop such a foundation for the design of the PISA 2012 context questionnaires.

This chapter is organised into five main sections:

- The first section ("General purpose and policy relevance of PISA") elaborates on the general purpose and policy goals of the PISA study and explains why, and what kind of, context information is needed to serve these goals. As a tool for comparative system monitoring, as a study that helps to understand the effectiveness of education policies and practices and as a database for policy-related research, PISA has to address various factors on the student/family, school and system level. Equity issues warrant special concern within education policy.

- The second section ("The general knowledge base: Research in educational effectiveness") establishes the general knowledge base that PISA can use to fulfil its functions. The section starts by establishing a rubric that helps to organise constructs and measures that operate at different levels. This taxonomy is based on research in educational effectiveness, which is subsequently outlined in some detail to justify the selection of input, process and outcome measures at the system, school, classroom and student levels. General findings on educational productivity and effectiveness are also presented here.
- The third section ("Learning conditions for mathematical literacy") deals with the study of learning environments in a way that is relevant to the assessment of mathematical literacy as the focal domain in PISA 2012. The questionnaires will be designed to add further outcome measures such as mathematical attitudes and strategies, indicators of instructional quality, including the possibility that students were exposed to the types of problems used in PISA (i.e. Opportunity to learn), as well as school- and system-level context variables.
- Grounded in this knowledge base and directed by the policy purpose of PISA, the fourth section ("Specifying the questionnaire design for PISA 2012") delineates the PISA 2012 questionnaire design. Three steps appear requisite: first, an overarching architecture for constructs and measures for future PISA cycles is established. Second, in order to put earlier work to appropriate use and to establish trend information, measures used in PISA 2003, which also focussed on mathematics as the major domain, are revisited and measures of proven quality and relevance are identified. Finally, implementation issues are discussed that are new to PISA: Applying a rotated design to the student questionnaire will allow for more material to be used in the study. Some questions will be answered by all students, as in previous cycles, some by sub-samples of students, while computer-based delivery of the school questionnaire will improve the user-friendliness and adaptability of questionnaire administration.
- The fifth section ("Assuring validity evidence, explanatory power, and policy relevance of the PISA design") examines principles of analysis and design that may help to keep PISA innovative and to ensure the validity of measures in the future, i.e. complex statistical modelling techniques to define and identify effects, assuring cross-cultural equivalence and validity of the measures, re-sampling schools that took part in PISA 2003 to capture change at the school level and using follow-up tests or longitudinal extensions to suitably address the added value of schooling.

Annex A includes the core and optional questionnaires implemented in the PISA 2012 cycle.

GENERAL PURPOSE AND POLICY RELEVANCE OF PISA

Key goals of PISA

PISA serves an array of parallel purposes. The views endorsed by different stakeholders in the participating countries and economies may be broken down into the following broad areas:

- PISA is a monitoring structure that provides reliable comparative information on education systems, describing system structures as well as the functioning and the productivity (i.e. the gross outcome or "yield") of education systems. PISA data cover student career paths up to secondary level, school characteristics, school governance, student performance and motivation as well as equity issues (such as performance by gender as well as socio-economic background).
- PISA is an international study contributing to our knowledge base on educational effectiveness. PISA observes patterns of relationships between inputs, processes and outcomes of education. Thus, it helps to understand how educational outcomes are produced. First, PISA allows for a decomposition of variation of student performance by the individual, school and system. Moreover, each PISA cycle provides data about multiple factors – covering these three levels – which, according to previous research, are expected to impact performance in reading, mathematics or science. In addition to describing these factors, PISA estimates their direct and indirect relationships to student performance and other outcomes. Large representative samples and sophisticated statistical models allow for the generalisation of findings both within and across countries.
- PISA provides a data source for the study of educational contexts in general (e.g. how family, school and out-of-school education interact in the development of life skills) and the study of educational variables in economic and sociological contexts (e.g. the relationship between demographics, economic wealth, economic growth and human resources). The database will become even more informative with the availability of trend data also for mathematics and will cover a dozen years as soon as PISA 2012 is included.

Thus, PISA offers three types of policy-relevant "products": *i*) Indicators that monitor the functioning, productivity and equity of education systems; *ii*) Knowledge on factors that determine educational effectiveness; *iii*) A reliable, sustainable, comparative database that allows researchers worldwide to study basic as well as policy-oriented questions.

Comparative system monitoring

A central goal of PISA is to monitor education systems in terms of student performance (literacy, or more generally, life skills), as well as non-cognitive outcomes (such as student motivation and well-being), educational careers, context variables (such as the students' cultural, ethnic and socio-economic backgrounds) and, finally, process characteristics at the school and system levels (including evaluation and accountability policies, student selection and allocation, parental involvement, staff co-operation and opportunities to learn). PISA also provides indicators that cover relationships between these factors, e.g. the so-called social gradient which measures the strength of the relationship between socio-economic status and performance, or the relationship between educational resources and outcomes. The most important use of PISA data concerns the development of input, process, outcome, and relational indicators that can supply the OECD education indicators programme. These indicators, in turn, trigger public debate, shape educational policy and inform decision making.

PISA provides an international perspective. The programme is designed to enrich national and regional efforts to gather detailed information on students, schools and education systems. It complements these sources by providing international benchmarks and examples of what is possible from around the globe. In a globalised economic and social world, these examples can provide new ideas and suggestive evidence for education policy at the national, regional or local level.

The policy relevance of this system monitoring enterprise is based on *i*) defining and operationalising cognitive and non-cognitive outcome measures that inform the selection and prioritisation of educational goals within participating countries; *ii*) examining and reporting factors that may be subject to control by policy and professional practice (so-called malleable factors); *iii*) providing international benchmarks that allow policy makers to ascertain what they may learn from other countries. The selection of indicators is generally guided by policy demands. Educational policy making must deal with the functioning of the school system (i.e. operational characteristics such as resources allocated to schools), with productivity (such as the gross level of student outcomes) and, last but not least, with equity (e.g. how resources are distributed).

Reporting policy relevant indicators requires not only the assessment of performance data such as students' mathematical literacy, but also data based on student, school and parent questionnaires covering a broad range of context, process and non-cognitive outcome measures. This document provides theory and policy arguments that guide the selection of constructs and the definition of questions and scales.

Prior to PISA, a number of quantitative and qualitative studies in comparative education provided insights into the history, the functioning and, to some extent, the effectiveness of education systems. However, PISA is unique in combining the following features:

- It provides rigorous data and an integrated set of indicators for the monitoring of education systems.
- It has a clear focus on yield, in terms of student performance measured at the end of compulsory schooling. It also intends to report on malleable features of education systems and institutions that may initiate policy decisions, intervention, and improvement.
- It provides trend indicators, thus allowing for the description of changes in a country's performance levels over time as well as the development of background variables, processes and non-cognitive outcomes and an examination of the relationships between these various constructs. The more PISA moves into repeated cycles of measurement, the more can be learnt from examining the stability and change of educational input, processes and outcomes, and their relationship over the years.

Multiple examples of indicators based on PISA context data can be found in recent editions of OECD *Education at a Glance* reports (OECD 2007a, 2008, 2009a, 2011, 2012a), such as:

- Relationship between immigrant background or socio-economic background and student performance (2007, Indicator A6; 2011 and 2012, Indicator A5).
- Profiles of top performing students, including their attitudes and motivation for science (2009, Indicators A4 and A5).
- Relationship between enjoyment of reading and reading performance (2011, Indicator A6).
- Relationships between resources and outcomes in education (2007 and 2008, Indicator B7), especially with regard to class size (2008, Indicator D2).
- Outcomes of vocational versus general educational programmes (2007 and 2008, Indicator C1).
- Use of evaluation and assessment in education systems (2008, Indicator D5).
- School quality from a parental perspective (2008, Indicator A6).

- Relationships between student background and access to (or motivation to participate in) higher education (2007, Indicators A4 and A7; 2008, Indicators A3 and A7).

One of the most important challenges to the development of the contextual framework and the resultant questionnaires is ensuring that indicators can be compared across cycles, while at the same time allowing for new indicators to be introduced and established. After a decade of international student assessment in PISA, with context questionnaires mostly related to the respective major domain of performance, the time appears ripe to structure and order all the constructs and instruments that may be used either in single cycles or across cycles as a source of trend information. In the future, the policy relevance of PISA depends on how well this challenge is met.

Understanding patterns of effectiveness in education systems

Indicators will direct public attention towards successful and less successful sectors of the education system, or to goals that have been met versus goals that still pose challenges. Thus, the main goal essentially is to guide priority setting and decision making in educational policy. Besides obtaining descriptions of strengths and challenges with regard to student performance and the conditions of teaching and schooling in their respective countries, policy makers also want to understand why students achieve certain levels of performance. In serving an important explanatory goal, PISA context instruments are designed to help to answer this question. Therefore, the PISA questionnaires must cover the most important inputs and processes of student learning at the individual, school and system levels. Statistical models, using these multi-level data, will help to understand the complex relationships of how these inputs and processes interact with student outcomes. If data on resources and costs are available, PISA may also help to understand efficiency, i.e. effectiveness in relation to investments.

During the first decade of PISA, each survey used variables that were specifically related to the respective main domain. For instance, reading experience, interest in reading, instruction in language classes, etc. were important factors in 2000 and 2009. In contrast, PISA 2003 included attitudes towards mathematics and mathematical activities during mathematics lessons, while PISA 2006 explored various types of science instruction and science-related beliefs and interest. The *index of disciplinary climate in the classroom* and the *index of teacher support in the classroom* were measured with regard to the test language instruction in 2000 and 2009, but with regard to mathematics instruction in 2003. The operationalisation and analysis of data were domain-specific, but the constructs, the hypothesised links to student outcomes and the interpretation of data were informed by general theories of educational effectiveness.

Using these domain specific context data, a number of insights have been gained, with obvious importance for teaching practice and policy. Within the context of the OECD, each cycle is completed with the publication of an initial report (OECD, 2001; OECD, 2004; OECD, 2007b; OECD, 2010). Each cycle, the OECD publishes a number of Thematic Reports; in-depth reports provide detailed analysis of the PISA data covering a wide range of subjects at the student, school and government levels (recent examples include: *Untapped Skills: Realising the Potential of Immigrant Students,* OECD, 2012b; *Public and Private Schools: How Management and Funding Relate to their Socio-economic Profile,* OECD, 2012c; and *Let's Read Them a Story! The Parent Factor in Education*, OECD, 2012d). Since 2012, the OECD also produces the PISA in Focus, a series of monthly education policy-oriented notes designed to concisely describe a topic of policy relevance.[1] Lastly, the OECD has also launched the series *Strong Performers and Successful Reformers in Education* which provides tailored policy advice to countries that approach the OECD, or show the successes of a country's strong-performing education system, enabling other countries to learn from those that do well.[2]

Beyond OECD publications, here are some interesting examples from the first PISA Research Conference that took place in September 2009.[3] While some examples are based on cross-national analyses others are based on PISA data from one country:

- Raising the general level of students' reading engagement represents a means of improving equity (Baye et al., 2009).
- Resource conditions of the school as well as the extent to which the school encourages students to use their full potential are both significant predictors of mathematics performance in Korea (Kaplan, 2009a).
- Instructional time is strongly linked to mathematical performance in Swiss schools, except for schools with advanced requirements (Angelone et al., 2009).
- In order to support students' development of scientific competency and their interest in science, a focused pattern of scientific study seems to be more successful than increasing the breadth of scientific activities (Kobarg et al., 2009).
- The extent to which schools prepare for careers in science is strongly associated with their students' wish to pursue a science-related career (Lie and Kjaernsli, 2009).

- Japanese students are interested in inquiry-based learning, whereas science teaching at the upper secondary level does not cater to that interest. This finding could contribute to an understanding of why Japanese students in PISA show relatively low levels of positive attitudes toward science (Yasushi, 2009).

Although the analysis of PISA data can make important contributions to the knowledge base for education policy and practice, there are limits that have to be taken into account. Most important is the fact that PISA is a yield study, assessing literacy and skills that have been accumulated over the lifespan, from early childhood through different levels of schooling until the age of 15 years. PISA does not ascertain how much learning has taken place in the secondary school at which a student is presently enrolled. Such an assessment would require that the student's performance level was ascertained at the time of entering his or her present school and compared with the same student's present performance. In so doing, one would obtain a measure of progress or "value-added" in performance associated with educational experiences in the particular school. However, we do not have measures of student performance prior to the ones measured in PISA at age 15. Teacher quality and its impact on student performance cannot be judged in PISA either – at least with the design that has been in place for over a decade. This is because a random sample of 15-year-olds is taken in each school rather than intact classes, thus precluding the collection of information regarding classroom-level instructional strategies. Finally, in one out of five countries that participated in PISA 2006, the majority of students had only recently been allocated to the schools in question, making it practically impossible to draw direct conclusions on school effects within these countries.

As Baker (2009) notes, the history of policy making informed by international comparative studies has seen a number of short-cut conclusions, based on too simple hypotheses as to the causes of performance differences at the system level. Also, econometricians have studied a number of issues in educational productivity, but still most of this work is descriptive in nature and does not allow for causal inferences (see Hanushek and Woessmann, 2010).

It is exceedingly difficult to draw causal inferences, such as concluding that a particular education policy or practice has a direct or indirect impact on student performance, based on an observational survey and assessment data of the kind collected in PISA (Gustafsson, 2007; Kaplan, 2009b). If, for example, links between high student performance on the one hand, and school evaluation data being accessible to the public (as a school level policy) on the other hand were found, the design of the study would not warrant causal interpretation. This is because data on at least some potentially important factors, such as prior student performance, cannot be collected in PISA. As a consequence, such potentially important factors cannot be included in the analyses and therefore not controlled statistically. Thus, PISA cannot tell if this policy happens to be applied in high achieving schools, or whether the policy actually results in higher student performance. The fundamental problem is that, in the absence of random assignment to a treatment (a policy or practice), it remains generally unknown whether unobserved factors exist that are related to selection of the treatment and to the outcome of interest. The drawing of causal inferences thus relies on the researcher's willingness to make additional, often un-testable, assumptions. Some researchers may attempt to match control students to the treated students on some observed control variables and discard students for whom there are no suitable matches. The problem still remains, however, that selection may be driven by unobserved factors. In any event, it is essential that researchers both clearly state all assumptions made and, ideally, assess the sensitivity of their causal inference to the violations of these assumptions. (See the section "Assuring validity evidence, explanatory power, and policy relevance of the PISA design" for a discussion of advanced analytical methods and designs intended to address some of these issues.)

Much of the value of the programme is based on a constant interplay between PISA as a monitoring survey and more rigorous kinds of effectiveness research done elsewhere. Correlational and other exploratory results from PISA may subsequently be tested in longitudinal, experimental or intervention studies. Even so, factors that have been demonstrated to be relevant for educational effectiveness or efficiency in the research literature are prime candidates for continuous monitoring within PISA and for incorporation into the OECD system of educational indicators (see the section "The general knowledge base: Research in educational effectiveness").

Building a sustainable database for policy-relevant research

The influence of PISA already reaches far beyond educational practice, policy and research. PISA data are increasingly used also by economists and social scientists to examine broader issues such as the impact of human capital on economic growth (Hanushek and Woessmann, 2009) or how to predict successful integration of migrant families (Stanat and Christensen, 2006).

Broadening the scope of PISA as a database for policy-relevant research requires that general constructs such as student socio-economic status and immigrant background are operationalised in a highly sophisticated way. It also implies that

additional information and data may be required (e.g. on civic values and health). When conceptualising the content of PISA questionnaires, it should be clear that the database that PISA generates will meet demands – both in research and in policy decision making – that go beyond the system of indicators that has been established thus far.

In the long term, one of the major benefits of the PISA database will be the availability of trend data. Educational production processes can hardly be understood from cross-sectional data, because it is practically impossible to include all relevant variables – including information on previous performance – within the design. However, changes in performance at the country level (cf. Gustafsson, 2007; Hanushek and Woessmann, 2010) or even changes at the school (see the section "Extending the sample to study change at the school level [international option]") and the individual student level (see the section "Longitudinal extensions at the student level [recommended for future cycles]") can be interpreted and explained substantively, once changes in input and process are available. This goal requires that PISA defines a general set of variables that will remain constant over several cycles in the future (see the section "Distributing content across PISA cycles: An overarching design for general, domain-specific, and thematic extension variables").

Focus on equity

Equity refers to the fair distribution of educational provisions, learning opportunities and outcomes among participants in education with different characteristics (OECD, 2005a, p. 14). A first step in analysing equity is to study variation of cognitive and non-cognitive outcomes between and within schools. The gap between high achievers and low achievers is also of interest. Specific groups of students, defined by individual school or family characteristics can be compared. These are often based on geography (e.g. regions or urban/rural distinctions), public/private school enrolment, socio-economic status, gender, immigrant status and the heritage language. Small outcome differences between such groups may indicate greater equity. Equity can further be assessed in terms of the distribution of access to schooling, learning resources and opportunities among these groups. Policies aiming at enhancing equity may either aim at an equal distribution of resources, or provide additional support for disadvantaged groups. Finally, a "meritocratic" notion of fairness accepts that diversity exists among students, but claims that differences in educational opportunities and outcomes should not be grounded in social origin (Cleary, 1968). PISA provides data that help evaluate education systems with respect to these different criteria for equity.

In addition to assessing the degree of equity, PISA 2012 can also be used to examine characteristics of systems and schools, including educational policies, that are likely to be related to equity. At the system level, such analysis may include comparisons of systems that differ regarding tracking and the funding and management of schools. Policies aimed at promoting equity encompass equalising resources or providing targeted support for schools with a disadvantaged student population, as well as regulations on remedial instruction, support with language learning and other extra-curricular activities. Similarly, at the school level, ability grouping practices with those to assist students with special learning needs appear to be relevant. Adaptive instruction as well as a supportive and well-structured learning environment can be related to equity at the classroom level. Moreover, the prevalence of such characteristics in more or less equitable systems can be compared.

PISA has included variables that are relevant for equity as essential variables since its first cycle was conducted in 2000. Thus, trends in equity can be analysed as well as associations with changes in education policies and aggregated school processes. Information may thus be obtained regarding changes in equity, both among and within countries, covering more than a decade.

THE GENERAL KNOWLEDGE BASE: RESEARCH IN EDUCATIONAL EFFECTIVENESS

A rubric of educational outcomes and predictive factors

School integrates young people from different backgrounds into society by allowing them to participate in academic work, social interaction, cultural activities, and general civic life. Altogether, schools fill multiple functions within modern societies, for example:

- teaching young people and preparing them for success in life, including academic and professional careers;
- providing grades and certificates to guide educational and professional careers and help align adolescents to career pathways that fit their profile of competencies and personal goals; and
- the school as an institution with rules and norms – some explicit, some informal – educates minds and hearts.

PISA defines literacy as the capacity of students to use texts of various kinds, mathematical tools and science-based reasoning as they encounter, interpret and solve problems and make decisions in real-life situations. In modern, knowledge-rich societies literacy is one of the most important goals and outcomes of schooling. PISA also attempts to measure broader cognitive skills such as strategies for learning and thinking and problem solving competencies. Success in school – and in life – additionally depends on being committed, sharing values and beliefs, respecting and understanding others, being motivated to learn and to collaborate, and being able to regulate one's own learning behaviour. These constructs can be perceived as prerequisites of cognitive learning, but may also themselves be judged as educational goals, as the OECD publication *Defining and Selecting Key Competencies* (DeSeCo) has elaborated (Rychen and Salganik, 2003). Therefore, PISA addresses non-cognitive outcomes like attitudes, beliefs, motivation, expectations and learning-related behaviour, such as self-regulation, strategies and invested time. As one such non-cognitive outcome, truancy has received increased attention as an important (negative) indicator of students' use of learning opportunities, which is also predictive of school drop-out and deviant behaviour (Kearney, 2008; Lee and Burkam, 2003). These non-cognitive outcomes are measured mainly within the student questionnaire but also in the school questionnaire. International options, such as the educational career or the information, communication and technology familiarity questionnaire, extend those measures into areas of particular interest for a number of countries.

A large section of the student questionnaire, the school questionnaire and the international options, however, is devoted to contextual factors that are linked to cognitive and/or non-cognitive outcomes. These are used to understand effectiveness and to define indicators. The factors can roughly be classified as being either input or processes. Input factors are mostly related to the individual's social and personal background. Also, structural features like school size and funding are treated as inputs. Processes include learning and teaching as core processes with various variables designed to capture their quantity and quality. Moreover, professional activities by teachers and principals, as well as school policies and practices, are classified as process variables.

A particular challenge with international surveys is that contextual factors affecting student learning occur at four different levels, i.e. student/family, classroom, school and country. The basic production process goes from:

- Background and structural factors (inputs) through interactions, activities and policies (processes) to performance and non-cognitive outcomes: for instance, students' mathematical performance supposedly depends on the profile of mathematics-related activities in the classroom, which again may depend on teacher competency, class size, availability of technology and other resources.
- Higher-order levels (system, school) to classroom to student level processes and outcomes: the issue of whether the decision making authority is at the school or at some higher administrative level will partially determine instructional leadership and teacher collaboration, which shapes teaching strategies, and finally students' learning activities.

The basic structure of the Input-Process-Outcome model was developed in the 1960s for the International Association for the Evaluation of Educational Achievement (IEA) (Purves, 1987). A recent version of this model, as shown in Table 6.1, accommodates these dimensions. The first column displays four levels: students, classrooms, schools and countries. The three production phases are then given in the remaining columns, i.e. inputs, processes and outcomes, respectively. Each cell contains several example constructs. The rubric is quite comprehensive: it includes most of the factors that have previously been studied in PISA and practically all constructs are covered in this framework. Although it is a selection from the broad set of constructs discussed in the educational effectiveness literature, the table is still too comprehensive from a research design perspective. The number of constructs that will finally be able to be covered depends on the time available for the questionnaire and the rotation design applied.

Some input factors are fairly stable and difficult to change while others can be shaped by school development activities or policy decisions. Processes are usually more malleable, at least indirectly (e.g. by teacher education and professional development) and outcomes reflect the effects of the inputs and processes. Nevertheless, the discrimination between the three strands of variables is by no means clear-cut. Outcomes from one educational setting become input for the next, while some process aspects (e.g. learning strategies) may well be treated as input or outcome, depending on a given theoretical perspective, research design or practical considerations.

When turned into a dynamic model of school effectiveness (see Creemers and Kyriakides, 2008), outcomes become inputs for further development. Mathematics anxiety, for example, can be an outcome of schooling as well as an input, impacting, for instance, upon students' homework activities. Moreover, inputs may have reciprocal effects upon each other. For example, a school's socio-economic composition in many education systems is correlated with funding, parental involvement or even teacher quality. Each of these factors, in turn, allows for other (better) teaching-learning environments to be offered, which attract students (or, rather, parents) from higher socio-economic backgrounds, so that finally social stratification, resources and process quality are mixed and are difficult to disentangle.

In order to understand these relationships and effects correctly and to build sophisticated and adequately specified analytical models, we have to refer to knowledge accumulated in educational research. Two areas of research are especially important: research on educational effectiveness in general (discussed below) and domain-specific research on learning and teaching (see the section "Learning conditions for mathematical literacy").

Table 6.1 describes a constellation of factors at different levels that constitute an education system. Their interactions are complex and not fully understood and each level is important in shaping education policy and probable educational outcomes.

The main goal of educational effectiveness research is to identify "factors in teaching, curriculum, and learning environment at different levels such as the classroom, the school, and the above-school levels that can directly or indirectly explain the differences in the outcomes of students, taking into account background characteristics, such as ability, socio-economic status, and prior attainment" (Creemers and Kyriakides, 2008, p. 12). However, practices may neither be equally effective for all students within a school nor for all education systems, local contexts and schools. Moreover, depending on the kind of outcomes emphasised, different conclusions may be drawn (Kyriakides and Tsangaridou, 2004). Hence, modern research into educational effectiveness also takes interactions with input into account and examines differential effectiveness and adaptive practices. A considerable amount of research has been carried out in this field (e.g. Creemers and Kyriakides, 2008; Scheerens, 2000; Teddlie and Reynolds, 2000). The following sections will summarise results for each of the cells of the taxonomy introduced in Table 6.1.[4] Also, ways to operationalise major constructs and design considerations will be discussed at the end of each section.

Table 6.1
Two-dimensional taxonomy of educational outcomes and predictive factors

	Input	Processes	Outcomes
Students	Gender, grade level, socio-economic status Educational career, grades Immigration background, family environment and support ICT experience, attitudes, skills Openness, problem-solving styles	Attendance/truancy Outside-class activities - e.g. participation in after-school programmes Motivation, engagement Learning and thinking strategies, test taking strategies Learning time (including homework and private tuition)	Mathematical performance Mathematics-related attitudes, beliefs and motivation General school-related attitudes and behaviour, e.g. commitment, truancy Learning motivation, educational expectations
Classrooms	Class size, socio-economic background and ethnic composition Teacher education/training, expertise	Quality of instruction: structure, support, challenge Opportunity to learn: implemented curriculum, assigned tasks, mathematics-related activities Instructional time, grouping, assessment and feedback	Aggregated student variables
Schools	Socio-economic background and ethnic composition Affluence of the community School funding, public vs. private School size Parental involvement	Achievement orientation, shared norms, leadership, teacher morale and co-operation, professional development Admission and recruitment policies, tracking, course offerings/school curriculum, evaluation Teacher-student relations, supportive environment	Aggregated student variables Promotion/retention and graduation rates Attendance
Countries (Systems)	Economic wealth, social (in)equality Diversity policies	School funding, tracking and allocation, policies for professional teacher development, support for special needs and language minority students, hiring and certification policies Accountability and evaluation policies, locus of decision making	Aggregated student variables Average graduation level

System-level inputs and processes

In most countries, decisions about education policy and the allocation of resources are the responsibility of a national or federal entity. In more centralised systems, this may be a ministry with ultimate authority, while in other countries the functions may also be performed by federal state authorities or regional agencies at the district level. Although system level factors are organised in terms of input, processes and outcomes in this taxonomy, their role in analytical models most probably will be the role of moderating variables; i.e. system level factors have an impact on how input, processes and outcomes are related at the lower level.

Basic inputs at the system level concern the material resources and their distribution within a country. Studies have found that both within and across nations the level of per-capita income is a strong predictor of student performance (Baker et al., 2002). Affluence provides resources to facilitate educational performance (Baumert et al., 2005). At the same time, educational outcomes may also influence a country's economic well-being. Measures of societal inequality are linked to educational inequality in terms of family resources and learning conditions. Moreover, some countries provide for similar funding per student, while others specifically support students from rural, impoverished, or immigrant backgrounds, or those with learning or physical difficulties. A country's specific pattern of educational investment may have implications for both the level of performance as well as equity in outcomes and access to career pathways. Another important aspect that is closely related to societal inequality concerns policies on immigration and educational diversity (Stanat and Christensen, 2006). In a large number of countries immigrants perform at lower levels in some of the major subjects, but in other systems no differences are found or immigrants perform at a higher level than average (OECD, 2012b). Smaller performance gaps exist in countries that actively recruit highly educated immigrants. By contrast, a large proportion of immigrants with low socio-economic status may pose a special challenge to education systems. Research has further shown that language support policies successfully reduce disadvantages for students with a heritage language that is different from the language of instruction.

To describe school systems it is further important to examine policies regarding ability grouping and the responsibility for school management. Some countries allocate students to schools with different focuses: academic, vocational, professional or technical, according to their academic performance and/or their preferences. In such tracked systems, not all students are provided with the same learning opportunities. Therefore these practices have been subject to a recurring debate on whether educational separation denies academic opportunities to those students whose intellectual development is delayed as well as to those students of lower socio-economic status (Levin, 1978). Comprehensive systems, where all students are required to undertake at least a common core of academic work, may be more efficient in terms of equity (OECD, 2007b). However, even within schools academic classes may be heterogeneous in enrolments or streamed by ability. Besides their degree of tracking, countries also differ regarding the age at which students are allocated to different school forms as well as the possibility of changing tracks. Last but not least, school entrance and promotion policies will have an impact on PISA results; because – other things being equal – student performance will increase if students arrive at higher grades earlier (Gustafsson, 2007).

Previous PISA cycles have shown that public funding and management is most common, but in a majority of participating countries, a certain number of schools exist that are mainly funded and managed by non-governmental organisations like churches or businesses. Schools may further be privately managed but predominantly publicly funded; i.e. "government-dependent private schools". Although both kinds of schools exist within a government regulatory framework, private schools usually have considerably more freedom than public schools to make educational decisions. Thus they represent one form of decentralisation of decision making. Choice among public and private schools is considered as a mechanism to create incentives for schools to compete for students in order to improve the overall performance of the education system (Belfield and Levin, 2002). However, previous PISA results suggest that the higher average performance of privately managed and government-dependent privately managed schools is mainly due to the advantaged socio-economic backgrounds of their students and corresponding composition effects (OECD, 2007b). Hence, similar to the degree of tracking, the percentage of private schools and their characteristics may also affect equity (OECD, 2012c).

In recent years, many countries have been addressing the issue of how their schools can be more responsive to the needs of specific students, groups of students and communities. When decision making is vested in a central authority such as the national government, it is often argued that teacher hiring and training and the curriculum are too rigid to accommodate the needs of schools with different student inputs. For this reason, many nations have established a variety of reforms to decentralise their schools, placing educational decisions closer to the students and communities being served, in the hope that this will improve school performance (Hannaway and Carnoy, 1993). However, studies linking

this aspect to educational outcomes have not provided support for a strong association (e.g. Schmidt and McKnight, 1998). Some secondary analyses of international student assessments suggest that school autonomy is beneficial if combined with top-down evaluation or centralised exams (Woessmann, 2006).

Especially in decentralised systems it may be important to establish school and student assessments that will ensure that all students meet important educational standards and are equally prepared for their future educational careers (Carnoy et al., 2003). This idea is endorsed by several countries, but their approaches to evaluation and accountability still differ significantly. Some countries sponsor periodic systems of testing to gain standardised information on student and school performance at national, federal, or regional levels. Others require such evaluations, but leave it to the regional and local authorities to carry out testing and assessment, while a further group of countries lack systematic evaluations. Furthermore, accountability also extends to the consequences of evaluation results (Carnoy et al., 2003; Koretz, 2008). In some countries, tests and surveys only serve to inform the general public and educational entities. But test results can also be used to create incentives for higher performance, or to identify weaknesses and provide assistance to improve schools. Different systems may have different consequences for providing incentives to learn, uniformity of what is taught and educational results.

In addition to these rather descriptive variables at the system level, a number of more specific policies and practices have been proposed by the dynamic model of educational effectiveness (Creemers and Kyriakides, 2008). These aim at supporting schools realising a positive learning environment, as well as high quantity and quality of instruction and to offer students a large variety of learning opportunities. Few studies have integrated these system level processes. To a certain extent, PISA 2012 provides an opportunity to examine their prevalence and association with performance across a large sample of countries. Importantly, both the existence of these policies and the nature of their implementation might be analysed.

The learning opportunities offered are determined by national curricula. The IEA's Trends in Mathematics and Science Study (TIMSS) identified substantive variation in the intended and implemented curricula across countries, especially with regard to time invested, coherence and focus, which in part explains the observed differences in student performance at least for curriculum-oriented tests like TIMSS (Schmidt et al., 2001; Schmidt and Houang, 2007; see also the section "Opportunity to learn and quality of instruction: Assessing the learning environment"). Yet, in addition to formal learning, policy makers can also support and encourage the provision of extra-curricular learning opportunities. They may, for example, launch mathematics competitions or provide schools with resources and support for establishing after school programmes. Countries may also pursue different policies with regard to commercial tuition and other kinds of "shadow education" (Baker and LeTendre, 2005).

The quality of instruction can be controlled at the system level by defining standards for teaching and making sure that these standards are met. This implies the formulation of explicit expectations based on the latest research findings. Moreover, it presupposes that evaluations do not exclusively focus on measuring output but also examine teaching and instruction. By establishing respective policies, and organising and providing resources for in-service training, policy makers can further contribute to the professionalisation of school staff. Professional development is defined as "activities that develop an individual's skills, knowledge, expertise and other characteristics as a teacher" (OECD, 2009b, p. 49). In many countries, it is compulsory for teachers to spend a certain number of days on professional development. However, education systems vary significantly regarding the general level of participation and the pattern of attendance in different types of professional development. Participation may be encouraged by financial support, salary supplements and scheduled time. However, the OECD Teaching and Learning International Survey (TALIS) results suggest that the effects of these measures are limited (OECD, 2009b).

Finally, cultural norms and values also influence the behaviour of different stakeholders. Students' and parents' engagement in schools and learning partly depends on the general appreciation of education and related values. The overall status of teachers is important because it affects the perceived attractiveness of the profession for people considering taking up a career in teaching. Community involvement in schooling can provide greater sensitivity through local adaptation to the needs of local populations.

In terms of design considerations some of the system level variables, such as economic wealth, educational investments, equity of their distribution, composition of migrant populations, tracking and school entrance rules, may be captured from existing databases, e.g. from the OECD system of indicators on economy, welfare, and education.[5] Others could be covered by a system level questionnaire to be answered by senior administration officers, such as policies for teacher recruitment and professional development, evaluation and accountability, school funding, centralisation/decentralisation, standards and extra-curricular activities, support for special needs and minority students. Some

indicators may be derived from aggregated PISA data, like the index for school autonomy in decision-making that is derived from the school questionnaire. Student, teacher and parent data may be aggregated at the country level to provide measures for cultural contexts, such as appreciation of education relative to other personal and social assets.

School-level inputs and processes

The organisation of schools is complex and it varies considerably, not only from one country to the next, but also among educational sub-divisions within a country such as federal states or provinces, regions, school districts and individual schools. In discussing inputs at the school level, first of all the type of student input is likely to affect educational outcomes. Schools recruiting students of higher socio-economic status often tend to create general school environments with high expectations and support for academic learning. Moreover, such schools may attract better teachers. This provides advantages even for students of lower socio-economic status who are enrolled in these schools. The student input is likely to depend on the general affluence of the surrounding community which has additional effects on student learning. Wealthier neighbourhoods tend to be less preoccupied with the possible effects of crime and violence and they have more resources to support schools in informal ways. Reciprocally, good schools reflect and may attract engaged parents and thus also influence the community. A somewhat different dimension is the proportion of students with a heritage language different from the language of instruction. Linguistic diversity is, on the one hand, a resource for schools. On the other hand, it detracts from resources since specific support systems and adaptive teaching skills may be required to deal with comprehension difficulties and likely cultural differences. In many countries, the socio-economic context varies greatly across urban, suburban and rural communities.

When schools encourage and provide for parental involvement, parents can become more effective in supporting both school programmes and the educational progress of their children. Furthermore, parents who participate in school activities are more likely to volunteer their efforts in assisting the school, therefore increasing available resources. Parents who know what schools expect by being familiar with the school programme and teacher expectations are better able to assist their own children in learning. Finally, the act of attracting parents to participate actively in educational and school endeavours can serve to form social networks where parents get to know and help each other. Arguably, such social networks raise overall performance through the accumulation of "social capital" (Coleman, 1988).

School size is also thought to be linked to performance. With larger enrolment numbers, schools can offer their students a greater variety of teachers, courses, electives and extra-curricular activities. This enables students to select those courses that are most meaningful to them educationally and for which they are most motivated. Larger schools, however, also tend to be more impersonal and students may experience less individual support. Research has thus shown that smaller secondary schools demonstrate a greater and more equitable distribution of student engagement and performance among 15-year-olds (Coleman, 1988). Indeed, the question about optimal school sizes remains unanswered. In fact, some research suggests that school size affects different groups of students (e.g. those based on socio-economic status) in different ways (Lee and Smith, 1997).

In addition to these inputs, a number of processes directly or indirectly predict outcomes at the school level. Arguably, the most important malleable characteristic of effective schools is the school climate (or the environmental quality of the school). The school climate encompasses not only norms and values but also the quality of relationships and the general atmosphere. An academic focus – a general consensus about the mission of the school and the value of education – facilitates learning. In addition, an orderly learning atmosphere maximises the use of learning time. By contrast, disrespectfulness and an unruly environment are counterproductive for both teachers and students alike and distract from the actual mission of the school. An orderly environment might, in turn, be fostered by coherent, reliable rules, (e.g. dealing with student misconduct like absenteeism). Effective schools are able to react to the specific learning needs of students and the needs of the teaching staff. They provide remedial and enrichment classes for students with learning disabilities, for highly gifted students and for students with a heritage language different from the language of instruction. In addition, effective schools offer student counselling and homework assistance. More broadly, the variety and quality of extra-curricular programmes can also facilitate student learning and may be especially relevant for non-cognitive outcomes and students' future career paths.

Three school organisational aspects are relevant for improving teaching and the learning environment: collaboration among staff, professionalisation, and school leadership. Collaboration among staff provides practical and emotional support for teachers and contributes to their professionalisation. It encompasses different techniques that aim at

co-ordinating practice, making the curriculum more coherent, and providing mutual support and feedback. In many countries it is relatively common for teachers to exchange instructional material or to discuss the learning difficulties of individual students. More sophisticated forms of co-operation include collective learning activities such as observation visits, providing feedback, engaging in professional learning activities and joint activities across classes and age groups (OECD, 2009b). The latter practices in particular help to transform schools into learning organisations as well as providing constructive feedback for teachers and support for professional development activities specifically addressing teachers' needs. Principals or school heads need to deal with administrative tasks, such as legal and budgetary issues, facility management and public relations. However, the core of their work is to ensure high quality instruction and learning within their schools. During the last decades, the concept of instructional leadership, which puts special emphasis on this aspect, has received considerable attention in research literature (Blase and Blase, 1998; Hallinger and Heck, 1998; Heck and Hallinger, 2005; Krüger et al., 2007; Leithwood and Riehl, 2005; Witziers et al., 2003). However, empirical results regarding the relationships of school leadership with student performance are inconsistent, possibly because leadership can have a rather indirect effect that is mediated by other school and classroom-level processes.

The quantity of instruction can further be addressed by policies and regulations on lesson schedules, timetables, management of teaching time, student attendance and parental involvement. Guidelines on teaching, grouping procedures and teacher behaviour can help raise the quality of teaching. Opportunity to learn is strongly influenced by the curricula and the selection of instructional materials. However, the formal descriptions of educational operations in schools are often misleading because implementation differs so widely (Fullan, 1992; Fullan and Stiegelbauer, 1991). Accordingly, the PISA questionnaires attempt to obtain information on operations directly from the participants, and especially the school principal.

It is insufficient to simply introduce the aforementioned policies and practices as their actual effects should also be monitored. School evaluation can be used to identify strengths and weaknesses in the school's operations, monitor practices and assess their effects on outcomes. TALIS has shown that evaluation practices vary widely across countries, not only with regard to the frequency but also concerning the sources of information used, the focus of evaluations and the consequences enacted (OECD, 2009b). So far, little research examines the effects of school evaluations on outcomes (Creemers and Kyriakides, 2008).

A somewhat different characteristic of schools that impacts decision-making is the degree of centralisation (versus autonomy) of governance. In some societies, the recruitment of teachers, development of curricula and assessment of students is undertaken centrally. In other societies, the school is merely expected to follow directions set out for routine school operations. Still further, at the opposite end of the spectrum, the school is largely autonomous and is expected to choose its own teachers, to influence teacher compensation and to choose (and undertake) its own academic assessment of student performance. In recent years, school autonomy has been in focus as a way of overcoming bureaucratic rigidity and potentially impacting learning outcomes (Bottani and Favre, 2001; Chubb and Moe, 1990). However, results are not yet conclusive and decentralisation is still subject to debate.

In terms of design considerations thus far, most school characteristics have been addressed by asking the principal to comment on school resources (or the lack thereof), school curriculum (i.e. timetables, tracks, remedial and enrichment classes, extra-curricular activities), school climate (i.e. expectations, teacher and student morale, parental involvement, behavioural problems), and professional activities (i.e. teacher collaboration, shared norms, leadership, evaluation procedures). Adding data from a teacher questionnaire, aggregated at the school level, would allow for a broader and probably more valid perspective on these issues. This approach is possible in those countries that link PISA 2012 to the next cycle of TALIS (OECD, 2009b). In addition to information from the school questionnaire on these issues, student data, including information from the optional ICT Familiarity questionnaire, and parent data can be aggregated as indicators of student composition, social capital, school resources and school climate (e.g. availability of ICT, teacher-student relations, parental attitudes towards and involvement in the school) while the optional Educational Career questionnaire addresses, for example, career counselling at school.

In PISA 2012, the school questionnaire is also used to gather outcome data beyond the performance and conduct of those students who are directly tested. For instance, principals are asked to provide data on student truancy, promotion and graduation rates. Such "objective" behavioural outcomes may be more comparable at least within an education system than most attitudinal or self-report measures. Also, they directly represent school-level outcomes while student performance is largely predicted by classroom- and teacher-level variables.

Classroom-level inputs and processes

Most of a student's educational experiences at school occur in the classroom. Here, the student is exposed to subject content, curriculum materials, instructional strategies and to the specific composition of, and climate within, the class. As previous research has shown, proximal variables – among them classroom characteristics and practices – are more closely associated with student performance than more distal factors, such as school-level and system-level conditions (e.g. Wang et al., 1993).

The most important inputs at the classroom level are the classroom context and the teacher (Wayne and Youngs, 2003). The former encompasses the characteristics of fellow students, instructional groupings, and class size. The concern about the socio-economic and language background of students at the classroom level is similar to that at the school level. In many cases, this composition varies not only between schools, but also within schools, particularly where grouping into different classes occurs according to students' performance levels. Although the rationale for such grouping is to enable teachers to improve instruction by tailoring it specifically to student ability, there is considerable evidence that the educational impact of such grouping may widen the gap in performance between such groups (Gamoran, 1992; Oakes, 2005). Ability grouping within classes that is more flexible and occurs for certain tasks or certain periods of time only or heterogeneous ability grouping may, however, be useful to implement adaptive and co-operative classroom teaching practices (Slavin, 1990).

A specific aspect of classroom practice, which is arguably quite important in the context of assessments like PISA, is the everyday practice of student evaluation, assessment, and grading. Do students work on standardised tests regularly? Do they receive teacher feedback in terms of grades or written reports or oral feedback? Do schools, teachers, students and/or parents receive information about student performance as compared to general standards? How do they use this information? Do students practise test-taking? How do these practices evolve and change over time in different countries? As has been shown by research on formative and summative assessment practices (Brookhart, 2009), as well as in studies on conditions and effects of high-stakes testing (Borko et al., 2007; Koretz, 2008), student results in standardised assessments like PISA may interact with assessment-related policies and practices.

Class size is generally believed to have an (negative) effect on student performance. Smaller classes enable more opportunities to participate in class and facilitate the use of teaching practices that are targeted at individual (groups of) students. Moreover, fewer students mean that fewer assignments need to be evaluated leaving teachers more time for providing feedback and designing richer activities and assessments. And yet, international research suggests that class size reduction needs to be substantial before it has any effect, and it is less relevant in secondary schools compared to primary schools (Gustafsson, 2007). Moreover, effects of class size seem to be culture specific: comparatively large classes are found in many Asian countries which nevertheless realise a high level of average student performance.

The teacher is characterised by her or his education and training, teaching experience and expertise. Advanced academic degrees, a major in the subject being taught, and professional experience have been described as desired qualifications and in certain models as indicators of teacher quality. However, results regarding their association with student performance are inconsistent (e.g. Hanushek and Rivkin, 2007; Libman, 2009; Mullis and Martin, 2007; Zuzovsky, 2009). Stronger effects have been found for teacher expertise (e.g. Baumert et al., 2010; Hill, Rowan and Ball, 2005). This effect has been defined as the interplay of teachers' general pedagogical knowledge, content knowledge, pedagogical content knowledge, beliefs, motivation and capacity to self-regulate (e.g. Bromme, 1997; Brunner et al., 2006; Shulman, 1987). In addition to the inputs described above, a number of processes at the classroom level have been found to be relevant for educational effectiveness, particularly classroom climate and instructional quality. Research has shown that student learning is generally supported by a positive and respectful atmosphere that is relatively free of disruption and focused on student performance (Creemers and Kyriakides, 2008; Harris and Chrispeels, 2006; Hopkins, 2005; Scheerens and Bosker, 1997).

The major facets of a positive classroom climate are: supportive teacher-student interactions, good student-student relationships, achievement orientation, and an orderly learning atmosphere with clear disciplinary rules. Instructional quality, however, is a more complex aspect. Existing evidence suggests there is no single best way of teaching. Well-structured lessons with close monitoring, adequate pacing and classroom management, clarity of presentation and informative and encouraging feedback (i.e. the key aspects of "direct instruction") are linked positively to student performance. These components help create an orderly classroom environment and maximise effective learning time. Yet student motivation and non-cognitive outcomes benefit from additional characteristics of instructional quality, such as a classroom climate and teacher-student relations which support student autonomy, competency and social relatedness. Furthermore, in order to foster conceptual understanding, instruction has to use challenging content (Brown,

1994). Also, different student sub-populations may benefit from different instructional practices. Thus, teachers have to orchestrate learning activities in a way that serves the needs of their specific class. Klieme et al. (2009) condensed this knowledge into a framework of three "basic dimensions of instructional quality": *i*) clear, well-structured classroom management, *ii*) supportive, student-oriented classroom climate, and *iii*) cognitive activation with challenging content. Several independent studies of secondary school mathematics education have now confirmed this triarchic structure of classroom quality and given some support for the cognitive and motivational impact that was hypothesised (*TIMSS-Video*: Klieme et al., 2001; *COACTIV*: Baumert et al., 2010; *Pythagoras*: Lipowsky et al., 2009). Klieme and Rakoczy (2003, see also Kunter et al., 2008) identified similar structures within national extensions to PISA. However, instructional quality, especially cognitive activation, is to a large extent domain specific. Aspects specifically related to mathematics learning will be discussed in the section "Learning conditions for mathematical literacy".

In terms of design considerations, with the exception of a few countries that extend their samples to include grade-based components, the PISA samples do not include a classroom level as 15-year-old students are randomly sampled from all classes in a school. This sampling design presents challenges to the examination and classroom and teacher effects (Opdenakker and van Damme, 2000; Van Landeghem et al., 2005) and ultimately the examination of well-specified models of educational effectiveness. In addition, it reinforces the yield nature of the PISA assessment. Still, PISA can report on classroom inputs and processes in a descriptive way, taking up variables that have elsewhere been shown to relate to student performance and assessing them by means of student, school, and (where possible) teacher questionnaires. For example, principals have responded to questions concerning teacher background and predominant pedagogical orientations while students have been questioned about classroom context and practices. The PISA scales Disciplinary Climate in the Mathematics Classroom and Teacher Support in the Mathematics Classroom are indicative for the structure and support dimension in the triarchic model of instructional quality, respectively. PISA 2006 added several indicators for advanced, challenging practices in science education, and PISA 2012 will do so for mathematics (see the section "Opportunity to learn and quality of instruction: Assessing the learning environment"). Undoubtedly, details, for example, concerning teachers' constructivist vs. direct transmission beliefs and professional development activities would further add to this line of inquiry. For PISA 2012, this is likely to be achieved by some countries linking PISA 2012 schools to the sample for the TALIS teacher survey in 2013. In this way, such information, aggregated at the school level, would be available for secondary analyses of PISA data.

Student-level inputs and processes

By 15 years of age, it is not only the accumulated effects of schooling that contribute to an individual's academic performance, but also the experiences encountered at home. More educated parents are able to provide a richer set of learning opportunities. They are also able to provide more access to written materials for reading, travel, and other resources that engage their child's curiosity. Research has shown that parents holding high expectations for students' academic performance and showing interest in students' school work is linked to the educational success of their children, as are parents' participation in school conferences and involvement in homework (Alexander et al., 2007; Christenson, 2004; Hoover-Dempsey and Sandler, 1997; Ma, 1999; Sui-Chu and Willms, 1996; Wang et al., 1993).

One of the main purposes of gathering data on the family background is to take into account these influences on learning when estimating school effects. In addition to the student questionnaire, the parent questionnaire as an international option collects information on several relevant variables such as career expectations, parental involvement in school, and discussion of school related matters at home. Collectively, these instruments allow for the triangulation of parental support and academic expectations from varied perspectives. The ICT questionnaire provides data on computer availability and computer use – including at home – related skills and attitudes, which may especially help to explain the results of computer-based tests of cognitive skills.

In addition to socio-economic and family background, the linguistic background of students appears relevant. Growing up bilingually may foster the development of language awareness and facilitate the learning of additional foreign languages (Hesse et al., 2008). However, the possibly poorer competency in the language of instruction during the first years of schooling can also negatively affect learning in other subjects and in the long run be disadvantageous for students' educational pathways (Schmid, 2001; Stanat and Christensen, 2006). On the one hand, immigrants and ethnic minorities may have to adjust to an unfamiliar cultural context at school (Berry, 1980, 1990; Hovey and King, 1996; Liebkind, 1996) and they may face discrimination (Amiot and Bourhis, 2005; Perreault and Bourhis, 1999). On the other hand, quite often these groups hold stronger aspirations and higher motivation than peers from ethnic majority families (Kao and Tienda, 1998; Krahn and Taylor, 2005; Stanat, 2006; Stanat and Christensen, 2006). PISA strives to identify the various inputs, processes and outcomes of migrant education in order to disentangle migration effects from

socio-economic and language effects. Also, it seeks to separate individual level from compositional or contextual effects. PISA 2012 strives to extend the explanatory power of the study by including measures of acculturation, and perceived cultural proximity, as well as in-depth information on language background within the optional educational career questionnaire. The optional parent questionnaire provides additional information on migration background while the optional educational career questionnaire provides additional information on linguistic background and upbringing.

General student-level processes mainly refer to the learning time. While formal learning opportunities are usually compulsory, students may still limit their use, e.g. by skipping classes, by arriving late or by being inattentive during lessons. Absenteeism and time spent on task significantly influence student performance and they are related with school drop-out rates, delinquency and drug abuse (e.g. Baker et al., 2001; Lee and Burkam, 2003; McCluskey et al., 2004; Wilmers et al., 2002). Moreover, these behaviours are also relevant as an aggregate indicator of the school-level learning environment. By contrast, students may extend their learning time and the variety of learning opportunities by individual study or by participating in extracurricular activities like academic clubs, competitions, volunteering and debating. Students with a disadvantaged social background seem to especially benefit from participation in such programmes (e.g. McComb and Scott-Little, 2003). In some countries, taking private tuition and commercial courses, like Juku in Japan, is a common and important part of the education system (Baker and LeTendre, 2005).

Educational aspirations are a significant predictor of the students' future educational attainment (e.g. Thiessen, 2007), and they vary between different ethnic groups regarding their quality and stability (e.g. Kao and Tienda, 1998; Mau and Heim Bikos, 2000). Student motivation and attitudes towards learning are important student inputs, which may also be interpreted as non-cognitive outcomes. Just like students' learning styles, these are to a large extent subject specific and will therefore be discussed in the next section "Learning conditions for mathematical literacy". General, domain-independent student characteristics, in contrast, can be used to explain student performance in general, especially cross-curricular competencies such as problem solving. PISA 2012 will cover a measure of students' openness for learning and exploration, plus a taxonomy of problem-solving approaches, and a measure of test-taking strategies.

In terms of design considerations, while family background and support would best be captured by the parent questionnaire, most learner characteristics have been and will continue to be assessed within the student questionnaire (especially since the parent questionnaire is an international option in 2012). Two of the international options proposed for PISA 2012 provide additional information, namely the educational career and the ICT questionnaires. The strongest student-level predictors of learning (i.e. intellectual capacity and previous domain knowledge), however, cannot easily be measured in PISA. On the other hand, longitudinal extensions to PISA, which have been successfully applied in some countries, have been able to take predictors such as intellectual composition and prior knowledge into account to appropriately estimate the value-added measures of school effects. Consideration should be given to longitudinal extensions as an option if policy makers want to be informed accurately of the effects of schools and schooling (see the conclusions in the section "Assuring validity evidence, explanatory power, and policy relevance of the PISA design").

LEARNING CONDITIONS FOR MATHEMATICAL LITERACY

PISA 2012 focuses on mathematics as the major domain of student assessment, with clear implications for the design of both student and institutional level background variables. In line with the educational taxonomy shown in Table 6.1 at the student level, PISA aims to portray important aspects of the affective domain, information about students' experience with mathematics in and out of school (e.g. experience of different approaches to teaching mathematics, preferred ways of learning), motivation, interest in mathematics and engagement with mathematics. At the institutional level (i.e. classrooms and schools in Table 6.1), PISA aims to portray important aspects of learning and instruction in mathematics, including an investigation of the relationship between learning and teaching strategies and performance as well as the relationship between school organisation, structures and active student engagement with learning. These aims, in turn, are in line with the "Longer-term strategy for the development of PISA", in which the PISA Governing Board as early as 2005 stated that "it would be conceivable to make the effectiveness and efficiency of educational processes the overarching theme for the PISA 2012 assessment", because such an analysis "would lend itself particularly well to the subject area of mathematics, as the most school-bound subject covered by PISA".[6]

In the following sections, foundations for the study of various facets of mathematical literacy, including individual prerequisites and institutional conditions will be specified. Moreover, conceptual issues will be supplemented by recommendations for actual measures to support the development for the PISA 2012 context questionnaires.

Mathematical literacy: A challenge for instruction and assessment

Since mathematical competency is one of the most important prerequisites for success in the modern, technology-rich and knowledge-driven world, it is given priority status by most education systems. Mathematical competencies have been and still are an essential domain in nearly every large scale educational assessment – be it national or international. High-quality mathematics instruction has received strong attention by educational policy makers worldwide. However, the very meaning of high quality instruction has changed dramatically over the last twenty years (see Schoenfeld, 2006). This reform movement originates from many sources, from Dewey's pragmatism, the late Russian psychologist Vygotsky's notion of cognitive development and German "Reformpädagogik" (i.e. reform pedagogy) to modern constructivist epistemology and the theory of self-regulated learning. Based on these sources, non-mechanistic concepts of learning and teaching were invented, which eventually permeated the professional practices of mathematics teachers.

Probably the most influential document of recent years that has shaped the understanding of mathematical education and has been echoed in many parts of the world, the "Curriculum and evaluation standards for school mathematics" was issued in 1989 by a professional teacher organisation in the United States. This document sets out "five general goals for all students: *i*) that they learn to value mathematics; *ii*) that they become confident in their ability to do mathematics; *iii*) that they become mathematical problem solvers; *iv*) that they learn to communicate mathematically; and *v*) that they learn to reason mathematically" (NCTM, 1989, p. 5). The NCTM standards mark a substantive shift in the way mathematics education is conceptualised, because general competencies for mathematical thinking substitute the content-oriented learning goals that had previously characterised the purposes of education. Mathematics education now aims at fostering mathematical thinking and its application to real-world problems, well beyond declarative knowledge and procedural skills. Subsequent research elaborated and expanded on those competencies through both theoretical work and empirical studies (Niss, 2003; Blomhoj and Jensen, 2007). Most recently, similar developments have been further expressed through the Common Core State Standards Initiative (CCSSI) (2010) for Mathematics in the United States. As well as standards for mathematical content, the CCSSI draft document includes a set of standards for mathematical practice that in part build explicitly on the NCTM standards mentioned above.

The mathematics assessment framework for PISA has been built on this new tradition, and in fact for one decade, PISA has been a major force in promoting the thinking/modelling approach to mathematics education, strongly supported by policy makers as well as mathematics education experts worldwide. The PISA 2012 Assessment Framework for mathematics (Chapter 1 of this report) does not just mirror the traditional content strands of school mathematics, such as algebra, geometry or statistics. Rather, it stresses the "big ideas" that guide conceptual understanding, and it requires mathematical competencies well beyond technical knowledge and skill.

When PISA was conceptualised, the fundamental idea was to assess mathematical thinking in context. In the draft framework for PISA 2012 mathematical competence is defined as "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" (see Chapter 1). When essentially the same concept was elaborated and transformed into test items for earlier PISA surveys, the items were designed as "measures of the extent to which students presented with (problems mainly set in real-world situations) can activate their mathematical knowledge and competencies to solve such problems successfully" (OECD, 2004, p. 37). The assessment framework for PISA 2012 (Chapter 1) continues to base the assessment of mathematical competence for 15-year-olds on phases of the mathematical modelling process, and builds on a central set of mathematical capabilities, namely *communication, mathematising, representation, reasoning and argument, devising strategies, using symbols and formal operations,* and *using mathematical tools*. Also, items are developed in line with four broad content categories of secondary school mathematics, namely *space and shape*, *change and relationships*, *quantity* and *uncertainty and data*. This means the PISA mathematics survey for the foreseeable future will continue to go well beyond technical mathematical content knowledge and skills.

The PISA cognitive assessment thus covers three out of the five goals of modern mathematics education cited above, namely students' ability for mathematical reasoning, problem solving and communication. Both remaining goals, valuing mathematics and being confident in doing mathematics, are assessed as non-cognitive outcomes within the PISA student questionnaire, along with mathematics-related strategies. Besides being warranted as outcomes of mathematics education, these constructs will also assist in explaining differences in performance on the PISA mathematics test. The next section will provide the conceptual background and measures.

Outcome measures related to mathematical literacy: Strategies, beliefs and motivation

Strategies and metacognition

Mathematics educators want to know not only which and how many tasks or problems students can solve, but also *how* they are working on these problems (i.e. which strategies they apply in studying mathematics, preparing for tests, or tackling/approaching problems). Domain-specific strategies of learning and problem solving are both an important outcome and a prerequisite for mathematical learning – much like text-based learning strategies within reading education. Good knowledge and command of strategies allows for the transfer of mathematical literacy into new contexts and new tasks.

In the context of reading literacy, PISA has applied student self-assessment of three cognitive learning strategies, namely memorisation (e.g. learning key terms, repeated learning of material), elaboration (e.g. making connections to related areas, thinking about alternative solutions), and control (i.e. meta-cognitive strategies that involve planning, monitoring, and regulation). PISA 2003 adapted these scales to the domain of mathematics. The elaboration scale used several items that had been adopted from the domain of reading, like "When learning mathematics, I try to relate the work to things I have learnt in other subjects". In addition, the elaboration scale included specific items like "When I am solving mathematics problems I think of new ways of getting the answer" and "… I think of how the solution might be applied to other questions". This modified elaboration scale proved to be a valid and strong predictor of mathematics performance across countries. At the same time, memorisation does not seem to be identifiable as a separate and valid scale at all, and mixed results have been found for self-assessed control strategies (OECD, 2005b, p. 297; Vieluf et al., 2009a, 2009b).

It is not by chance that the mathematics-specific elaboration items have been situated in the context of solving mathematics problems. International comparative research has shown that mathematics instruction all over the world is mostly defined by working on problems (Christiansen and Walther, 1986; Hiebert et al., 2003; Stigler and Hiebert, 1999). When students practise individually, in small groups, or as a whole class, they are regularly assigned either to problems or routine tasks. Even when new content is introduced, teachers in many countries use example problems to work with the class through new solution methods. Thus, studying (for) mathematics or learning mathematics basically means solving assigned problems. Rather than learning strategies, as in the domain of reading, problem solving strategies are at the heart of mathematical literacy. Cognitive research in mathematics learning re-specified these as instantiations of what is now commonly called metacognition (Desoete and Veenman, 2006; Garofalo and Lester, 1985; Schneider and Artelt, forthcoming; Schoenfeld, 1992). In general, metacognition refers to the individual's own awareness and consideration of his or her cognitive processes and strategies, as defined in the now classical paper by Flavell (1979).

Design considerations: to allow for trend analysis, the elaboration strategies scale from PISA 2003 (a self-assessment of mathematical learning and problem-solving strategies) was retained. However, as Schneider and Artelt (in press) point out, self-assessments of knowledge about and/or use of strategies are weak indicators of metacognition. In addition, indicators based on logfile data from the computer-based cognitive skills tests, or indicators of declarative metacognitive knowledge similar to the PISA 2009 instrument which assessed metacognition in reading could be considered in the future.

Motivation and intentions

In an era when many countries suffer from low enrolment in Science, Technology, Engineering, and Mathematics (STEM) subjects, especially among females, the strengthening mathematics interest and motivation or more general positive attitudes towards mathematics has become a major policy issue. Also, evidence abounds of a positive relationship between mathematics-related attitudes and mathematics performance (see, e.g. the meta-analysis by Ma and Kishor, 1997).

Theories of student motivation usually distinguish between intrinsic and extrinsic motivation. Intrinsic motivation comes from rewards inherent to a task or activity itself. PISA 2003 assessed this type of motivation with regard to mathematics by means of a scale labelled, interest in and enjoyment of mathematics (example item: "I am interested in the things I learn in mathematics"). The research literature suggests that intrinsic motivation has an effect on student engagement and time-on-task, learning activities, performance and career choices, and it can be shaped by classroom instruction (Kunter, 2005; Rakoczy et al., 2008; Ryan and Deci, 2000). Extrinsic motivation, by contrast, comes from outside the subject, i.e. from expected rewards. It has been found to be an important predictor of course selection, career choice and performance (Wigfield et al., 1998). In PISA 2003, a scale for instrumental motivation to learn mathematics was applied. One of the items reads as follows: "Learning mathematics is worthwhile for me because it will improve my career options".

More sophisticated models of behaviour regulation distinguish between more general dispositions like intrinsic and extrinsic motivation, and actual intentions. Intentions are closer to real decision making, and are more relevant for assessing how students deal with mathematics in their present life and in their plans for the future. Such items are presumably less biased by culture-specific response styles or peer-group norms. Short-term intentions may be assessed by a scale with items like "I will try to work hard to make sure I learn mathematics" (Lipnevich et al., forthcoming). Long-term intentions or future orientations may be captured by adapting a scale that was introduced for science-related studies in PISA 2006, with items like "I would like to work in a career involving mathematics". The latter scale would allow the examination of whether education systems have a chance to recruit high achieving students for career paths which make use of these competencies. Also, issues of equity in access to STEM careers can be investigated.

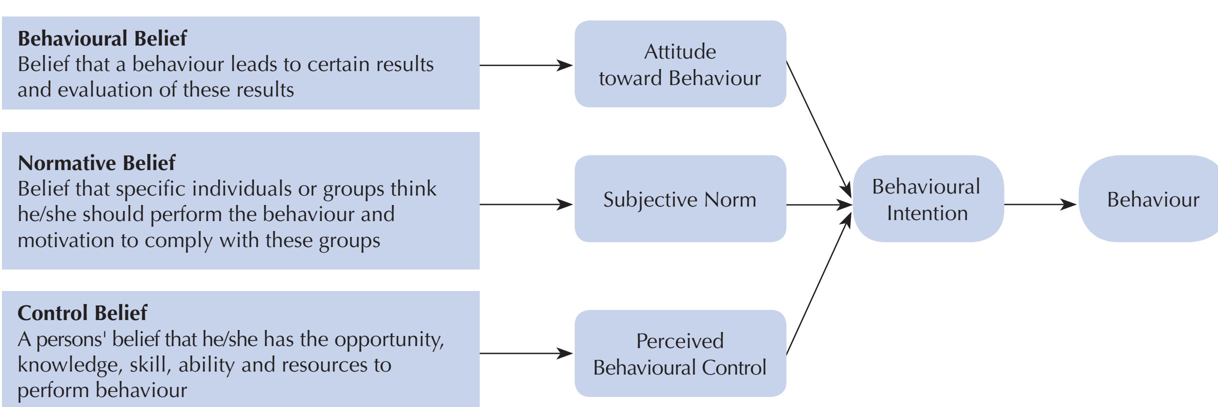
In terms of design considerations, these four scales, namely interest in and enjoyment of mathematics, instrumental motivation to learn mathematics, short-term intentions and long-term intentions will suffice for addressing policy-relevant issues in student motivation for mathematics learning. Applying six attitudinal scales within the student questionnaire plus several embedded measures within the cognitive test, as in PISA 2006, would unnecessarily overburden students. However, innovative approaches were tried out to reduce cultural bias and assure the cross-cultural comparability of results (see the section "Extending the sample to study change at the school level [international option]").

Self-related beliefs and planned behaviour

How students think and feel about themselves is an important predictor for how they act and decide when challenged by tasks and situations (Bandura, 1977). While self-efficacy is defined as one's conviction or belief about his or her capability to cope with certain tasks, self-concept is defined in a broader sense as the overall perception of one's personal attributes based on continuous self-evaluation. Research into mathematics education found that subject-specific self-efficacy predicts mathematics grades, mathematical problem solving, interests, and – once again – career choices (Britner and Pajares, 2006; Pajares and Miller 1997; Turner et al., 2004). PISA 2003 assessed both mathematics-related self-concept (example item: "I learn mathematics quickly") and self-efficacy in relation to mathematical tasks (see Lee, 2009; OECD, 2005b). More specifically, students were confronted with eight tasks, ranging in difficulty from straightforward technical procedures like solving the equation $3x + 5 = 17$ to mathematical modelling tasks like calculating the petrol consumption rate of a car (OECD, 2005b, p. 292), and asked "How confident do you feel about having to do the following calculations?". Both self-concept and self-efficacy turned out to be cross-culturally equivalent, strong predictors of student performance (Vieluf et al., 2009a, 2009b).

How confident students are about their ability to solve mathematical tasks, as well as how students value mathematics, are highly relevant factors in predicting or explaining student behaviour with regard to mathematics, e.g. course-taking and career decisions. A number of expectancy value models both in psychology and in economics have been proposed to integrate both aspects of decision-making. One such model is Ajzen's (1991) theory of planned behaviour, which states that volitional behaviour is determined by specific attitudes and subjective norms (i.e. value component) plus perceived behavioural control (i.e. expectancy component). (For an application to mathematics in cross-cultural research, see Lipnevich et al., forthcoming.) According to Ajzen (1991), by manipulating these predictors, the chance that the person will intend to do a desired action can be increased. This, in turn, will increase the likelihood of the behaviour actually occurring (see Figure 6.1).

■ Figure 6.1 ■
The theory of planned behaviour



In the current context, students' attitudes, perceptions of control and subjective norms, may predict their desire to spend time on mathematics homework, to ask questions in class, or to engage in relaxation exercises to counteract anxiety, and thus enhance their mathematics performance.

Design considerations: application of the full Ajzen model substantially increases the analytical power of the study for explaining mathematics-related effort, mathematics-related student behaviour, and indirectly, if possible, student outcomes. However, because of the larger costs and efforts that are needed for measuring all facets of the model, this extension is given to a subset of students only, based on the suggested rotation design (see the section "Rotation design for extended student questionnaires and computer-based delivery"). Within the student questionnaire, self-concept and self-efficacy should be used again and a measure of mathematics anxiety retained. However, as its cross-cultural validity and incremental predictive power are questionable (Lee, 2009; Vieluf et al., 2009a, 2009b), new item formats were tried out for this scale.

Opportunity to learn and quality of instruction: Assessing the learning environment

PISA intends to cover learning conditions for the major domain in a particular cycle for two reasons: first, PISA helps to describe and compare learning environments for 15-year-old students in different countries, based on measures that have been shown to be important indicators of quality instruction elsewhere in the research literature. Second, those measures can help to explain student outcomes – not in every country, but in those countries where most students have been attending the school in which testing occurs for at least two years – which, for example, is the case in two thirds of countries that participated in PISA 2006.

Learning environments that help students to acquire mathematical literacy – as defined in PISA – have been described in the research literature as learning in context (Schoenfeld, 2006), discussion-based teaching (Boaler and Greeno, 2000; Stein et al., 2008), and they include mathematical modelling (Blum and Leiss, 2005) as an essential activity. In previous, seminal work, Schoenfeld (1987, 1992) called for a mathematics culture that allows for meaningful learning, makes connections between mathematical concepts, and links these constructs to students' everyday life – as compared with traditional instruction that aims at mastering formulas, algorithms, definitions, and other technical content. The "current vision of effective mathematics teaching", as Stein et al. (2008) refer to it, includes the following: "Students are presented with more realistic and complex mathematical problems, use each other as resources for working through those problems, and then share their strategies and solutions in whole-class discussions that are orchestrated by the teacher" (p. 315).

As PISA intends to cover major aspects of the learning environment for mathematics within its questionnaire design, this vision of modern mathematical education becomes a challenge. In order to describe teaching and learning processes that are conducive for mathematical literacy, it is insufficient to ask about content coverage. First, because mathematical literacy as defined by PISA is a yield measure, dependent on fundamental mathematical abilities and processes, and hence less dependent on specific content elements that have been taught at school. Second, the quality of mathematics instruction is not determined by content but rather, by clarity and structure, support for students, and cognitive challenge (see the section "Classroom-level inputs and processes"). Therefore, PISA needs to re-define the notion of opportunity to learn (OTL) that has been an important construct characterising learning environments in other international student assessments. In addition, within the PISA design, OTL data obtained from students are examined at the school level.

Opportunity to learn – Coverage of content categories and problem types

The notion of opportunity to learn was introduced by John Carroll in the early 1960s, and was initially meant to indicate whether students had sufficient time and received adequate instruction to learn (Carroll, 1963; Abedi et al., 2006). It has since been an important concept in international student assessments (Husén, 1967; Schmidt and McKnight, 1995; Schmidt et al., 2001), and has been shown to be strongly related to student performance, especially in cross-country comparisons (Schmidt and Maier, 2009, pp. 552-556). At the same time, the meaning of the construct became much broader. Stevens (1993, pp. 233-234) already identified four kinds of OTL variables most prevalent in research:

- Content coverage variables: these variables measure whether or not students cover the curriculum for a particular grade level or subject matter.
- Content exposure variables: these variables take into consideration the time allowed for and devoted to instruction (time-on-task) and the depth of the teaching provided.
- Content emphasis variables: these are variables that influence which topics within the curriculum are selected for emphasis and which students are selected to receive instruction emphasising lower-order skills (i.e. rote memorisation) or higher-order skills (i.e. critical problem solving).

- Quality of instructional delivery variables: these variables reveal how classroom teaching practices (i.e. presentation of lessons) affect students' academic performance.

Thus, for certain authors, OTL has become more or less a synonym for the quality of instruction experienced by the student. Schmidt and Maier (2009), however, in their review argue that opportunity to learn (OTL) is a rather uncomplicated concept: "What students learn in school is related to what is taught" (p. 541), and they intentionally focus on OTL "in the narrowest sense: Student's content exposure" (p. 542).

Schmidt and Maier acknowledge that although OTL may be a straightforward construct, it is quite difficult to measure. In order to explain differences in the achieved curriculum, traditionally, teachers and/or students have been asked whether and how certain curricular content has been realised in instruction (the implemented curriculum), sometimes using logs (Rowan et al., 2004). In addition, curriculum experts have been asked whether and how content elements have been covered within curricular documents like syllabuses, textbooks and standards (the intended curriculum). From these raw data, various indicators have been extracted: in many cases, the content taught has been judged twofold, in terms of topic and level of demand, while at the system level indices for coherence, rigor and focus have been derived (Schmidt and Maier, 2009).

Design considerations: in PISA, the measurement of OTL has to be modified from approaches used in other studies as the mathematics assessment is not framed according to content elements, but refers to seven fundamental mathematical abilities and four content categories. In PISA, the operationalisation of OTL is built mainly on student judgements.

PISA 2012 aims to identify country (and probably school) level profiles in learning opportunities. Students will be confronted with carefully crafted mathematics tasks – some representing mathematical abilities and content categories as mentioned in the PISA mathematics framework, some representing more traditional tasks asking for procedural and declarative knowledge. Following each of those items, students are asked to judge whether and how often they have seen similar tasks in their mathematics lessons and in previous assessments. Thus, it is possible – aggregated at the country, but possibly also at the school level – to measure learning opportunities in a way that allows for differentiation between types of problems and content.

In addition, students will be asked to estimate their familiarity with certain mathematical concepts. This measure, as well as the self-efficacy scale described in the section "Outcome measures related to mathematical literacy: Strategies, beliefs and motivation", may be used as proximal indicators for opportunity to learn.

Learning time

Following Carroll (1963), numerous studies have shown that learning time is a major predictor of student outcomes in many subjects, including mathematics. PISA 2012 will apply a measure that was already used in 2003, to ensure trend analysis. Students were asked: "On average, how many hours do you spend each week on the following?" with items for homework, remedial and enrichment classes, work with a tutor, and out of school classes. In addition, the weekly number and duration of mathematics lessons will be gauged in both the student and the school questionnaires.

Quality of instruction

As mentioned in the section "Classroom-level inputs and processes", research on secondary school mathematics instruction suggests that *i*) structure and efficient classroom management; *ii*) teacher support and student orientation; and *iii*) challenge and cognitive activation have to be addressed as basic dimensions of instructional quality. Two of these dimensions are covered by existing PISA 2003 scales, namely *disciplinary climate in the mathematics classroom* and *teacher support in the mathematics classroom*, both in the student questionnaire.

The international PISA 2003 report (OECD, 2004) showed that the construct *disciplinary climate in the mathematics classroom* was strongly associated with mathematical literacy, while other variables – such as class size, mathematical activities offered at school level, and avoidance of ability grouping – had no substantial effect once the socio-economic status was taken into account. These findings are in line with previous research and with the theory outlined in the section, "Classroom-level inputs and processes" which predicts that well-structured, efficient classroom management is a prerequisite for student learning. However, the PISA 2003 report did not study the relationship between learning environment and affective outcomes, such as students' interest and enjoyment in mathematics. Meanwhile, Vieluf et al. (2009a) have shown that teacher support is positively linked to students' interest in mathematics, also after controlling for socio-economic status.

Challenge and cognitive activation, however, are very difficult to assess, although this dimension is crucial for fostering mathematical literacy. Similar concepts, emphasising particularly the importance of a demanding orchestration of teaching the mathematical subject matter which gives learners vast opportunities to develop and practice competencies, are broadly discussed in mathematics education (e.g. Blum and Leiss, 2007). However, none of the indicators used in PISA 2003 could capture this dimension. Several approaches have been used successfully in national studies and are proposed for the PISA 2012 field trial:

- As homework is an almost universal element of mathematics instruction, the way teachers deal with homework during lessons can be used as an indicator for challenge. Rakoczy et al., (2005) developed a scale capturing the process-oriented dealing with homework, with sample items "My mathematics teacher is eager to learn how we solved problems in our homework assignment" and "When talking about homework results, we try to understand and correct student errors". (Note that the scale does not address homework assignments per se, or time used for homework, but the quality of classroom discourse about homework results).
- Students can also be asked about the type of items they usually encounter when practicing in their mathematics classes, e.g. "We usually have to think for a while in order to solve the problems we are assigned by our mathematics teacher". (Scale "cognitive activating tasks" by Baumert et al., 2008, used in PISA 2003/Germany.)

Teaching practices and students' mathematical activities

A behaviour-oriented version of the triarchic model of instructional quality has been implemented in the OECD TALIS study by asking teachers how often they implemented each of 13 given practices in their teaching:

- Structuring practices (five items): e.g. "I explicitly state learning goals." Other items include a summary of former lessons, homework review, checking the exercise book and checking student understanding during classroom talk by questioning students.
- Student-oriented practices (four items): e.g. "Students work in small groups to come up with a joint solution to a problem or task." Other items include ability grouping, student self-evaluation and student participation in classroom planning.
- Enhanced activities (four items): e.g. "Students work on projects that require at least one week to complete." Other items include making a product, writing an essay and debating arguments.

Based on TALIS main study data from 23 countries, it has been shown that *i*) the three dimensions can be differentiated across countries (i.e. the triarchic model has some cross-cultural validity), *ii*) structuring practices, as hypothesised, are associated with higher levels of classroom discipline (as perceived by teachers), and *iii*) participation in professional development as well as teaching high-ability classes raises the frequency of using these practices. Mathematics and science teachers report less student orientation and less frequent use of enhanced activities than did teachers who teach other subjects (OECD, 2009b).

In terms of design considerations, the TALIS teaching practices scales were adapted for use in PISA 2012 within a brief add-on to the TALIS teacher questionnaire. Items for enhanced activities would have to be reframed for mathematics, with examples such as: "Using mathematics to solve an everyday problem", "Connecting mathematics to other subjects", "Inducing rules from mathematical patterns", "Elaborating a proof of a mathematical theorem", "Discussing examples and counterexamples", "Using multiple graphical representations for a mathematical problem", "Comparing different solution methods, or multiple solutions for the same task."

While a link to TALIS will be feasible in only some PISA countries, all of the students can be asked how often they actively participate in such activities.

School- and system-level support for teaching and learning of mathematical literacy

Most of the school- and system-level variables that have been discussed in the section "System-level inputs and processes" and in the section "School-level inputs and processes" are likely to be linked to mathematics performance and/or mathematics-related attitudes. Input and process indicators specific to mathematics can be obtained by aggregating student responses about OTL and quality of instruction (see the section "Opportunity to learn and quality of instruction: Assessing the learning environment") to higher levels.

However, a number of mathematics-specific variables at the system or the school level will directly shape conditions for teaching and learning mathematical literacy. Therefore, special emphasis may be given, at the system level and at the school level.

At the system level to reform:

- initiatives, school improvement activities and professional development programmes targeted at mathematics;
- the role of mathematics in student admission, tracking, assessment and promotion as well as in school evaluation; and
- the intended instruction time for mathematics per year.

At the school level to reform:

- mathematics-related course offerings (i.e. remedial lessons, enrichment activities, tutoring sessions);
- qualification of the mathematics teaching staff; and
- collaboration among mathematics teaching staff.

In terms of design considerations, given that PISA is asking the school principal about all of these issues, the data may also be aggregated from the school to the system level.

SPECIFYING THE QUESTIONNAIRE DESIGN FOR PISA 2012

Distributing content across PISA cycles: An overarching design for general, domain-specific and thematic extension variables

Some of the relevant factors in understanding student performance, attitudes, and behaviours, and the functioning of education systems are straightforward (such as demographic variables, previous educational career choices, instructional time and class size), some have been well established in previous PISA cycles (such as student socio-economic status, cognitive strategies and school-level decision-making), while others have proven to be less easily addressed within the PISA design (e.g. accountability policies at the system level, teacher variables, aspects of the classroom learning environment or out-of-school activities). Choosing among the many variables that might be incorporated into the design is a complex process, directed by the priorities that countries have set for the study, but also informed by educational research, as outlined in the previous sections.

Given that PISA has multiple purposes, and addresses multiple outcomes of student learning, and considering also that trend information is widely required, the context questionnaires must cover:

i) General variables (for all cycles)

- General input variables:
  - Student-level inputs (grade, gender, socio-economic background = parental education and occupation/family wealth/educational resources/cultural possessions, migration data = immigration status/heritage language/age on arrival in country, family support)
  - School-level contexts and inputs (community size, resources, qualifications of teaching staff)
- General process variables:
  - School-level processes (decision-making, admission policies, assessment and evaluation policies, professional development, teacher engagement/morale, teacher-student relations, parental involvement)
  - Instructional processes (learning time, disciplinary climate, teacher support)
- General outcome variables:
  - General non-cognitive outcomes – Commitment to learning (behavioural: truancy; personal goal: educational expectations; motivational: learning engagement; affective: sense of belonging)

ii) Domain-specific trend variables (for major domain only, included every nine years)

- Domain-specific non-cognitive outcome variables (strategies and metacognition, domain-related beliefs, self-related beliefs, motivation)
- Domain-specific processes variables (opportunity to learn, instructional quality, system- and school-level support)

iii) Thematic extension variables (extensions within individual cycles)

- International options (e.g. in 2012, educational career; ICT familiarity; parent questionnaire)
- Context variables for additional domains (e.g. ICT-related experiences relevant for computer-based problem solving)

- Descriptive and explanatory variables for specific reports (e.g. in 2012: mathematics-related motivations and intentions)
- Malleable variables at the school level (e.g. tracking policies, teacher certification) that are specifically selected for descriptive purposes or for causal inference

iv) System-level data, gained from INES or system-level questionnaire

- Output of educational institutions (e.g. certificates)
- Financial and human resources invested into education
- Access to and participation in education
- Learning environment and organisation of schools

In previous PISA cycles – especially in PISA 2006 – the questionnaire design largely neglected constructs that were not related to the major domain. Thus, the set of general variables *i*) was sometimes restricted to include student inputs only. Student, and to some extent also school, questionnaires focused on processes and non-cognitive outcomes that were domain-specific. However, this design is inappropriate for measuring trends.

Finding an appropriate balance between *i*), *ii*), *iii*), and *iv*) is crucial for the overarching design of PISA questionnaires, and for the long-term success of the PISA programme.

In order to establish valid and reliable trends at the country level, it is important to implement a constant set of general variables in all cycles both for the calculation of proficiency estimates and as major reporting variables. Thus, these context and input background variables should not change.

Policy makers do not only need trend information on student performance. Education systems serve other goals as well. Thus, policy makers ask for information on domain-independent, non-cognitive or behavioural outcomes, namely educational aspirations (occupation expected at age 30), school commitment (sense of belonging), truancy and learning motivation (e.g. effort and perseverance) – all covered in previous PISA cycles. So far, PISA authors have been reluctant to use these variables, fearing their insufficient cross-cultural comparability. However, as trends are available over an increasing number of cycles, the focus now is on change rates within countries, rather than on cross-sectional comparisons of status. For example, whether and to what extent school commitment is increasing or decreasing, will be a relevant indicator within countries. Also, cross-cultural validity will receive more attention (see the section "Cross-cultural validity").

Two process scales, namely *disciplinary climate in the classroom* and *teacher support in the classroom*, are related to student performance and student interest, respectively (see Klieme and Rakoczy, 2003; Vieluf et al., 2009a). These scales, in addition to learning time, as general process variables, are to be retained in future cycles. Instead of changing the wording each time, however (i.e. instead of applying them to the major domain only), the wording should be kept stable and domain-independent, in order to allow for change models to be applied.

Depending on the major domain of cognitive assessment, additional measures shall be taken into account (domain-specific trend variables). And finally, each cycle will have its specific foci, either in the form of international options, or as e+xtensions to compulsory questionnaires.

Exploration of measures from PISA 2003

The following aspects of the learning environment for mathematical literacy were captured in the 2003 data set:

- Class size (student questionnaire)
- Disciplinary climate in the mathematics classroom (student questionnaire)
- Teacher support in the mathematics classroom (student questionnaire)
- Time spent on mathematics instruction at school and at home (student questionnaire)
- Frequency of ability grouping for mathematics (school questionnaire)
- School activities that promote student engagement with mathematics, such as participation in mathematics competitions (school questionnaire)
- Characteristics of the mathematics teaching staff, namely level of qualification, teacher consensus and evaluation policies (school questionnaire)

Several mathematics-related outcome variables were assessed as well: learning strategies, interest in and enjoyment of mathematics, instrumental motivation, self-efficacy, self-concept and mathematics anxiety.

In order to deliver trend information on non-cognitive outcomes and mathematics-related context/process variables, PISA 2012 retained as many variables that were used in 2003 as possible, unless they were shown not to work cross-culturally or not to account for differences in outcomes. Table 6.2 gives an overview of all questions from the PISA 2003 student and school questionnaires. The first row identifies the individual inputs – student demographic and family background variables – which have remained stable over the cycles.[7] These input variables are used to study the distribution of educational resources and outcomes within countries, and they are treated as control variables in analytical models.

The second row of the table contains school input variables such as the public versus private status of the school and its financial, staff, and other resources, including the size and composition of the student population. These variables are inputs at the school level but budgetary decisions, overall changes in the programme structure, and privatisation policies are at least partly controlled at the country or system level.

Based on the technical characteristics documented in the PISA 2003 Technical Report (OECD, 2005b), taking also into account multilevel analyses of predictive power (Vieluf et al., 2009b), cross-cultural validity (Vieluf et al., 2009b) and a review of the use of information from the school questionnaire in previous PISA cycles (Hersbach and Lietz, 2010), a set of variables that will be retained as school input variables has been selected.

At the individual student level, results of the predictive power analyses have shown that the link between learning preferences and mathematics performance were relatively weak; the link for co-operative learning was particularly weak. In similar analyses of PISA 2000 data (Jehanghir and Glas, 2007), the relationship to reading performance in PISA 2000 of co-operative learning was shown to be non-significant and that of competitive learning to be rather small. Hence, these two scales might be deleted. Among the three strategy scales, memorisation as a domain-specific learning strategy was not related to mathematics performance. In addition, cross-cultural invariance – even at the metric level – could not be established for this scale. In contrast, elaboration was strongly linked to performance and was acceptable in terms of cross-cultural invariance (see the section "Outcome measures related to mathematical literacy: Strategies, beliefs and motivation").

At the school level, relatively strong effects were found for *sense of belonging, time spent on mathematics homework* and *disciplinary climate in the mathematics classroom* as process predictors of mathematics performance as well as mathematics interest. Consequently, these variables were retained whereby disciplinary climate in the mathematics classroom demonstrates the strongest (scalar) level of invariance, which allows for comparison of mean scores between countries.

Self-efficacy and *self-concept*, both these motivation scales have scalar invariance, allowing use of these variables as non-cognitive, mathematics-related outcome indicators, establishing trend, as well as cross-national comparative analyses. However, alternative measures were tried out to further improve cross-cultural comparability.

Rotation design for extended student questionnaires and computer-based delivery

Covering all of the general, domain-specific trend and thematic extension variables would definitely go beyond the scope of time and space allocated for context questionnaires in PISA. This potential problem especially holds true for the student questionnaire, which is used, by means of data aggregation, to generate information for all levels of the study design (i.e. from the individual student up to the system level). The school questionnaire is also in danger of being extended, facing an increased risk of non-responses. To allow PISA 2012, as well as future cycles, to serve the intended multiple purposes, and to prevent an increase in missing data, new methods to reduce response burden are needed. To this end, two technical innovations were introduced in PISA 2012.

Rotation

PISA 2012 introduced a rotation design for the student questionnaire, similar to the design for cognitive items, which makes use of item packages distributed over a number of different booklets. Each student is allocated one of these booklets, and thus receives a limited number of cognitive tasks, while the booklets taken together cover a large "universe" of items from various domains. Booklets are randomly rotated among test-takers within each test site (i.e. school), hence the name of the design. Such rotation designs are common in large-scale assessments, both in national and international surveys.

Applying a rotated design to questionnaires is relatively new to educational measurement, though it is a well-established technique in other kinds of surveys. The implications of such a design for *i*) cognitive proficiency estimates, *ii*) the international report and the reporting of trends, *iii*) further analyses, *iv*) documentation and structure of the international database, and *v*) logistics, have been discussed elsewhere (Berezner and Lietz, 2009). Taking the existing experience into account, and based on analyses of PISA 2006 data, a rotated design for the PISA 2012 student questionnaire was adopted. In this way, it is estimated that the amount of material covered increased by one-third.

Table 6.2
Classification of PISA 2003 questionnaires

	Student questionnaire	School questionnaire
Antecedents (Student background and school context)	**Age, Gender, Immigration background** **Heritage language** **Parental occupation, Parental education, Family wealth and possessions (= 4 indices)**, <u>Family structure</u> <u>Educational career</u>, **Current grade**, Study programme, <u>Reasons for attending</u>, Study time outside school, Expected educational level	**Community size**
School input (Controlled at the system/ school level, depending on governance structure)		**Enrolment per gender,** Proportion of 2nd language learners **Public/private ownership**, Grades covered, Schedules per programme **Quantity of teaching staff**, Availability of computers, Financial resources, **School resources (3 scales)**
General processes (Controlled at school level, sometimes at system level)	**Student-teacher relations**	Decision-making/different sorts of autonomy, Admission policies, Assessment policies, Options for 2nd language learners Teacher morale, Negative teacher behaviours, Teacher participation
Mathematics processes (Controlled at instruction or school level)	**Instructional time, Time spent on mathematics homework** **Class size** **Disciplinary climate, Teacher support**	**Ability grouping in mathematics** **Mathematics activities** **Mathematics teaching staff: Qualification, evaluation policies; Mathematics teacher consensus**
General non-cognitive outcomes	Attitudes towards school, Truancy, Sense of belonging to school	Negative student behaviours Student morale Grade repetition
Mathematics non-cognitive outcomes	**Interest in and enjoyment of mathematics, Instrumental motivation** **Self-efficacy, Self-concept, Mathematics anxiety** **Strategies** (control/ **elaboration**/ <u>memorisation</u>), <u>learning preferences: co-operative vs. competitive</u>	

Notes: <u>Underlined</u> constructs were to be deleted. **Bold** constructs were to be kept for PISA 2012. For those constructs that are neither bold nor underlined, there was no strong rational for deletion or inclusion.

The common, non-rotating part of the student questionnaire, should at least comprise the student-level general input variables (see the section "Distributing content across PISA cycles"). This is based on methodological and reporting requirements, theoretical work and analyses of 2006 PISA data (Berezner and Lietz, 2009). Table 6.3 lists the variables included in the common part.

Table 6.3
Student questionnaire: Common part

Question n°	Description
ST01	Grade
ST02	Country study programme
ST03	Age of student
ST04	Sex of student
ST05	Attend <ISCED 0>
ST06	Age at <ISCED 1>
ST07	Grade Repeating
ST08	Truancy - Times late for school
ST09	Truancy - Days unexcused absence
ST115	Truancy - Times skipped classes
ST11	Family structure
ST12	Mother's occupation (ISCO); Component of ESCS
ST13	Mother's educational level – Schooling (ISCED); Component of ESCS
ST14	Mother's educational level – Post school (ISCED); Component of ESCS
ST15	Mother's current job status; Component of ESCS
ST16	Father's occupation (ISCO); Component of ESCS
ST17	Father's educational level (ISCED) – Schooling (ISCED); Component of ESCS
ST18	Father's educational level (ISCED) – Post school (ISCED); Component of ESCS
ST19	Father's current job status; Component of ESCS
ST20	Immigrant background
ST21	Age of arrival in test country
ST25	Language spoken at home
ST26	General home possessions plus country-specific wealth items; Component of ESCS
ST27	Number of certain possessions in household; Component of ESCS
ST28	Books at home

In addition, the common part is suggested to contain constructs that are needed to explain differential effects on major minority groups. From a policy point of view, this includes all variables that may help to explain differences between students with and without immigration background.

Table 6.4
Student questionnaire: Rotated forms A, B and C

Form A		Form B		Form C	
Q. n°.	**Description**	**Q. n°.**	**Description**	**Q. n°.**	**Description**
ST01-28	Common Part (see Table 6.3)	ST01-28	Common Part (see Table 6.3)	ST01-28	Common Part (see Table 6.3)
ST29	Intrinsic and Instrumental Motivation for Mathematics	ST42	Mathematics Self-Concept (Q2, 4, 6, 7, 9); Mathematics Anxiety (Q1, 3, 5, 8, 10)	ST53	Learning Strategies (Control vs. Elaboration vs. Memorisation)
ST35	Subjective Norms	ST77	Teacher Support in Mathematics Class	ST55	Attendance of Out-of-School-Time Lessons
ST37	Mathematics Self-Efficacy	ST79	Teacher Behaviour: - Teacher Directed Instruction - Formative Assessment - Student Orientation	ST57	Total Hours of Out-of-School Study Time
ST43	Perceived Control of Mathematics Performance	ST80	Cognitive Activation in Mathematics Lessons	ST61	Experience with Applied Mathematics Tasks (Q1, 4, 6, 8); Experience with Pure Mathematics tasks (Q5, 7, 9)
ST44	Attributions to Failure in Mathematics	ST81	Disciplinary Climate	ST62	Familiarity with Mathematics Concepts
ST46	Mathematics Work Ethic	ST82	+ Anchoring Vignettes	ST69	Min in <Class Period>
ST48	Mathematics Intentions	ST83	Mathematics Teacher Support	ST70	N° of <Class Period> per Week
ST49	Mathematics Behaviour	ST84	+ Anchoring Vignettes	ST71	N° of All <Class Period> a Week
ST93	Perseverance	ST85	Mathematics Teacher's Classroom Management	ST72	Class Size
ST94	Openness for Problem Solving	ST86	Student-Teacher Relations	ST73	Experience with Word Problems
ST96	Problem-Solving Strategies (SJT)	ST87	Sense of Belonging to School	ST74	Experience with Procedural Tasks
ST101	Problem-Solving Strategies (SJT)	ST88	Attitude towards School: Learning Outcomes	ST75	Experience with Pure Mathematics Reasoning
ST104	Problem-Solving Strategies (SJT)	ST89	Attitude towards School: Learning Activities	ST76	Experience with Applied Mathematics Reasoning
ST53	Learning Strategies (Control vs. Elaboration vs. Memorisation)	ST91	Perceived Control of Success in School	ST42	Mathematics Self-Concept (Q2, 4, 6, 7, 9); Mathematics Anxiety (Q1, 3, 5, 8, 10)
ST55	Attendance of Out-of-School-Time Lessons	ST29	Intrinsic and Instrumental Motivation for Mathematics	ST77	Teacher Support in Mathematics Class
ST57	Total Hours of Out-of-School Study Time	ST35	Subjective Norms	ST79	Teacher Behaviour: - Eacher Directed Instruction - Formative Assessment - Student Orientation
ST61	Experience with Applied Mathematics Tasks (Q1, 4, 6, 8); Experience with Pure Mathematics tasks (Q5, 7, 9)	ST37	Mathematics Self-Efficacy	ST80	Cognitive Activation in Mathematics Lessons
ST62	Familiarity with Mathematics Concepts	ST43	Perceived Control of Mathematics Performance	ST81	Disciplinary Climate
ST69	Min in <Class Period>	ST44	Attributions to Failure in Mathematics	ST82	+ Anchoring Vignettes
ST70	N° of <Class Period> per Week	ST46	Mathematics Work Ethic	ST83	Mathematics Teacher Support
ST71	N° of All <Class Period> a Week	ST48	Mathematics Intentions	ST84	+ Anchoring Vignettes
ST72	Class Size	ST49	Mathematics Behaviour	ST85	Mathematics Teachers' Classroom Management
ST73	Experience with Word Problems	ST93	Perseverance	ST86	Student-Teacher Relations
ST74	Experience with Procedural Tasks	ST94	Openness for Problem Solving	ST87	Sense of Belonging to School
ST75	Experience with Pure Mathematics Reasoning	ST96	Problem-Solving Strategies (SJT)	ST88	Attitude towards School: Learning Outcomes
ST76	Experience with Applied Mathematics Reasoning	ST101	Problem-Solving Strategies (SJT)	ST89	Attitude towards School: Learning Activities
		ST104	Problem-Solving Strategies (SJT)	ST91	Perceived Control of Success in School

Computer-based delivery

Computer-based delivery allows for more flexibility, adaptivity, and efficiency in administering and scoring questionnaires. For example, filter questions can be easily applied to guide respondents through the questionnaire, inconsistencies can be checked on line, and no manual data entry has to occur. Thus, in the long run, computer-based delivery might also help to improve administration of the instruments.

PISA 2012 used computer-based delivery for the school questionnaire. Subsequent cycles may also adapt this technology for the student questionnaire and international questionnaire options.

ASSURING VALIDITY EVIDENCE, EXPLANATORY POWER AND POLICY RELEVANCE OF THE PISA DESIGN

This section presents several recommendations regarding the PISA study design and analysis plan, in order to enhance validity and cross-cultural comparability, to define value-added indicators at all levels, and to enhance the explanatory power of the study. Strengthening the design from a research methodology point of view will ultimately allow for a greater number, and increasingly sophisticated series, of policy-relevant insights to be supported by PISA data.

Current design and standard reporting methods allow, for example, the amount of remedial education offered by schools (as reported by school principals), and the amount of out-of-school tuition (as reported by students and/or parents) to be documented. Policy makers may ask: Do schools differ in the amount of remedial education offered? Do students from various social and ethnic backgrounds and various tracks (as identified in the PISA sampling framework) have equal access to remedial education? Is there positive discrimination in favour of disadvantaged students? Does private tuition compensate for lower opportunities for school-based remedial education? How do these relationships develop over time (across PISA cycles)? How does my country behave with regard to these figures compared to other countries?

More sophisticated modelling, however, is needed to answer the following policy questions, and more caution is needed in interpreting related findings: Does paid work prevent students from participating in remedial education and/or tutoring? Does individual participation raise student performance? Does offering remedial education raise mean performance at the school level, reduce variation in student performance within schools, or reduce the social gradient (i.e. the association between socio-economic status and performance) within schools? These are hypotheses about the impact of one treatment (i.e. paid work, tutoring and offering remedial education) on another variable (i.e. participation in these activities, student performance). Testing such hypotheses requires some kind of causal inference. The validity of these inferences depends on the availability of control variables that may explain selection and self-selection into these treatments (see the section "Understanding patterns of effectiveness in education systems"). Analyses should be undertaken country by country.

Allocation (or self-selection) for remedial activities is assumingly driven by recommendation or pressure from teachers and parents, and influenced by student grades, gender, socio-economic background and the educational programme the student attends. In order to control for these conditions and to cover treatment variables, all students should be questioned *i*) about the kind and duration of remedial mathematics they have received during the present school year; *ii*) about the grades that they received at the end of the previous school year; and *iii*) whether, at the end of the previous school year, teachers and/or parents recommended remedial training. Including these questions in the student questionnaire is a prerequisite for causal inference on this matter. Similar treatment effects could be studied in future cycles.

At the school level, the effect of a school's absenteeism policy and its implementation on student truancy, sense of belonging and performance should be tested. Among other aspects, every school principal could be questioned on whether absenteeism was perceived as a severe problem by school authorities three years ago, and what kind of measures schools have been put in place since that point in time.

How to model the effects of context variables on student learning and performance

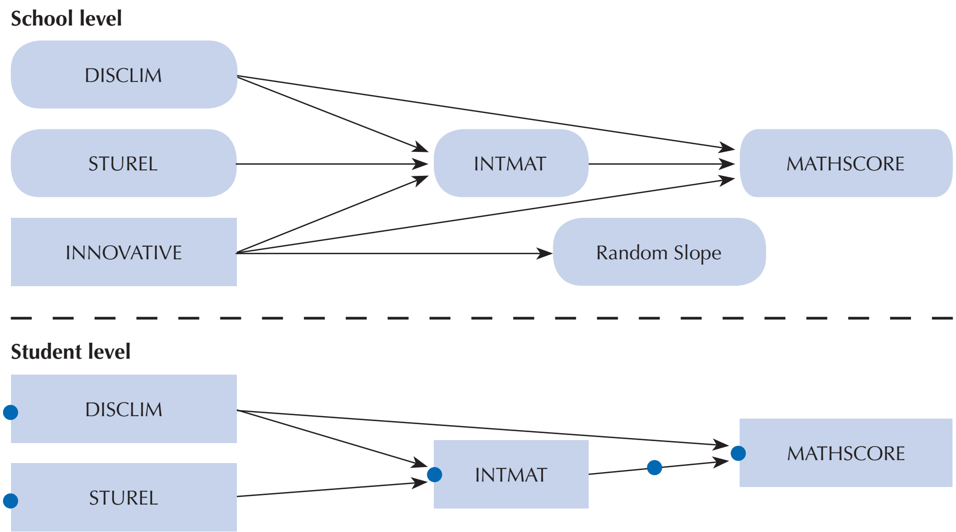
Previous PISA cycles and school effectiveness research more generally have studied the effects of school- and student-level processes on student performance and student motivation (e.g. interest in mathematics), controlling for inputs such as gender, socio-economic status and migration status (at the individual level) and the respective aggregates (at the school level). By including such control variables, erroneous "false positive" conclusions about relationships between two variables that are in fact only due to a third variable can be avoided. For example, in many countries, a large proportion of the difference in performance between immigrants and non-immigrants can be explained by the fact that immigrants often have lower socio-economic status and therefore less social and cultural capital. Ignoring the difference in socio-economic status would lead to an overestimation of the effect of the migration background. Moreover, PISA has used multilevel modelling techniques like hierarchical linear modelling (HLM) to account for the hierarchical structure of data with students nested within schools and schools nested within countries. An underestimation of standard errors that can result from a violation of the assumption of independent observations within the sample can thus be circumvented, because students within schools are likely to be more similar regarding several characteristics (e.g. Raudenbush and Bryk, 2002; Snijders and Bosker, 1999). Multilevel models further allow for the examination of effects at different levels; that is, the effects of school-level and individual-level input variables and processes.

However, research in educational effectiveness shows that – when modelling relationships between input, processes and outcomes – three additional observations from previous research should be accounted for namely that effects are often *i*) non-linear; *ii*) moderated by other variables; and *iii*) that they may be indirect, or mediated by third variables.

- Non-linear effects refer to the fact that more is not always better, but rather sometimes a medium value is optimal. For example, a study by Monk (1994) suggests that the relationship between teachers' subject knowledge and student performance may be curvilinear. On the one hand teachers are obviously not able to help students to understand content they are themselves not familiar with, but on the other hand, a teacher with very sophisticated knowledge may have difficulties communicating knowledge in a way that is easily understood by students.
- In line with the theory of differential effectiveness (e.g. Kyriakides and Tsangaridou, 2004), it is important to acknowledge that relationships between variables may not be similar in different sub-groups. For example, there is some evidence that students from diverse social backgrounds may benefit from different instructional techniques (e.g. Brophy, 1992; Walberg, 1986). Based on a constructivist understanding of student learning, current educational theory assumes that student learning is largely dependent on self-regulated processes, which are moderated by school, classroom and teacher factors. Modelling such differences requires the examination of interaction/moderation effects.
- Finally, it is reasonable to assume that not all effects on student outcomes are direct. Comparatively weaker effects on student outcomes are often found for policies at the school and system level, as compared to student background variables and classroom processes (e.g. Wang et al., 1993). This result may, in part, be due to the fact that the former variables do not exert a direct effect on students, but are rather related to school or classroom processes, which in turn have an effect on student performance. Moreover, school level variables such as school climate, shared values and norms, or procedures to deal with behavioural problems, may have a direct effect on non-cognitive outcomes (e.g. learning motivation, academic aspirations) and student behaviour (e.g. truancy, violence), while school effects on student performance and other subject-related outcomes (e.g. interest and self-efficacy beliefs) will most probably be mediated by teaching and learning within classrooms.

■ Figure 6.2 ■

An example of mediated moderation



Notes: DISCLIM – Disciplinary climate during mathematics lessons; STUREL – Student-teacher relations; INNOVATIVE – Innovative instructional strategies; INTMAT – Interest in and enjoyment of mathematics; MATHSCORE – Performance in mathematics. A dot represents a parameter estimated in the model.

In some cases *i*), *ii*) and *iii*) may apply simultaneously, for example when the effects of classroom climate and student-teacher relations on student performance are mediated by student interest. However, the effect of mathematics interest on student performance is stronger in schools where teachers use innovative, cognitively activating teaching practices,

than it is in schools with a more traditional approach to teaching. (This model is exemplified in Figure 6.2. It is partly based on results reported from a bi-national study of mathematics teaching by Klieme et al., 2009 and Lipowsky et al., 2009.) A non-linear effect may be included when it can be shown that neither a too lenient nor a too strict classroom climate is optimal for enhancing student interest or performance. (For a more complex example based on analysis of one country's PISA 2003 data see Kaplan 2009c; Kaplan et al., 2009.)

Moderation, mediation and non-linear effects, can be modelled with traditional HLM techniques. However, for modelling mediating processes the application of structural equation modelling (SEM) is more convenient (e.g. Kaplan, 2009c). The main advantage of multi-level SEM over multilevel regression analysis is that SEM uses latent instead of manifest variables, overcoming the biasing effects of measurement error at the student level (e.g. Muthén, 2002), but also for measures that are aggregated to the school level to model contextual effects (Lüdtke et al., 2008). Thus, using SEM models would improve the validity of conclusions about associations between different aspects of education systems. Using multilevel SEM also allows for more flexible combinations of mediation, moderation, and non-linear effects at different levels, as well as the analysis of categorical manifest and latent variables. For example, multiple-group SEM could be used to examine whether the effect of classroom disciplinary climate on student performance is similar across countries, or whether this variable may be more important in countries in which teachers have less authority. Latent class models could be applied for testing the hypothesis that it is a combination of different learning strategies rather than a high frequency of any one learning strategy that is related to student performance. Mixture models would allow the identification of groups of students within countries where relations between disciplinary climate and performance differ (e.g. Muthén, 2002). Such models may better reflect the complex reality of educational processes and their antecedents and consequences.

In short, using multilevel SEM models would help policy makers and professionals to gain a more comprehensive understanding of relationships among *i*) learning opportunities provided at the school level; *ii*) individual use of those opportunities (in terms of quantity and quality of individual learning activities); and *iii*) student outcomes, both cognitive and motivational. (e.g. Kaplan et al., 2009, based on the analysis of one country's PISA 2003 data.)

In summary, a simple production function approach to the study of educational outcomes is not appropriate from an educational point of view. System- and school-level as well as policy interventions will have an impact on student learning via processes of teaching, learning and collaboration. Researchers have to take these interrelationships into account when modelling system or school effects and drawing policy conclusions.

Cross-cultural validity

The aim of PISA is to compare student performance and its conditions across a large number of diverse countries. Cross-national studies do not only allow for analyses of policy approaches and other system-level inputs and processes and their relationships with performance, but also address issues relating to the transfer of findings from educational research to other cultures. Despite these benefits, cross-national studies also entail special methodological challenges. In addition to the regular tests of reliability and validity, cross-national research also needs to deal with the issue of equivalence and bias. Both refer to the question of time spent on tasks. Construct bias results from real differences in education systems, cultural norms and interpretations of various phenomena, but may also be a consequence of translation errors or method bias. To address this potential problem PISA involves experts in all of the participating countries. Some of the actual differences in the meaning of the constructs are handled through national adaptations, but finally SEM analysis is needed to understand the level of cross-cultural construct equivalence, and thereby the kind of between-country comparisons – i.e. comparison of means, profiles or correlations – that can be considered valid. For previous PISA cycles this issue has been addressed by using Multiple Group Confirmatory Factor Analysis (MGCFA) and examining whether the model has an acceptable fit across countries and whether the factor loadings are invariant. Results of these analyses show that PISA scales are quite equivalent across the whole sample of diverse cultures and education systems. For PISA 2012 the same analyses were repeated, both for the existing and for any newly developed scales. The latter were kept for the main study only when results were satisfactory.[8]

Item bias

Item bias results from individual items that behave differently from other items measuring the same trait for particular groups. This can be detected using Item Response Theory (IRT) based Differential Item Functioning (DIF) methods. In previous PISA cycles, DIF analyses have been carried out both in the field trial and in the main study. A "dodgy item report" was sent to country representatives, and problematic items were revised (in the field trial) or – if necessary – excluded. PISA 2012 used the same strategy.

Method bias

Method bias refers to a lack of comparability caused by the method used, i.e. by the sampling, by the administration or by the instrument itself. One instrument characteristic that may cause method bias is the use of Likert-type rating scales – a method that is frequently used in the PISA context questionnaires. Likert-type scales are especially susceptible to differences in response styles. Yet, culture is known to affect response styles. This interferes with the ability to examine mean scores for non-cognitive outcomes and to understand the relationships between contextual factors and performance (Hui and Triandis, 1985; Van de Vijver et al., 2008). Response styles may cause a lack of equivalence that can be detected with MGCFA or DIF analysis. However, as response styles often exert a similar influence on all items, additional analyses are required. A well-known problem in PISA is the difference in the correlations of certain scales with performance at the between-student within-country and at the between-country levels (e.g. Van de Gaer and Adams, 2010). For example, interest in mathematics correlates positively with mathematics performance within a country but the correlation at the between-country level is strongly negative. In other words, within countries higher-performing students report a higher level of interest in mathematics, as would be expected. At the between-country level, however, countries with higher mean performance have lower average levels of mathematics interest (e.g. Finland, Japan and Korea) whereas higher levels of mathematics interest are recorded for countries with lower mean performance (e.g. Brazil, Indonesia and Tunisia).

A more detailed analysis of method bias will help to preclude differences in response styles as a cause of differences in mean scores and correlations across countries. Three strategies were used in PISA 2012 to address this problem. First of all, analysis will be carried out to examine whether statistical adjustments are feasible for correcting the problem for at least some of the scales. An example is the use of a mixture modelling approach (Rost et al., 1997) to cluster countries into categories that are similar with respect to response style. The second approach is to examine the feasibility of addressing the cross-cultural validity problem using a wide variety of new item types and other item manipulations, such as those suggested by Buckley (2009). These include reverse keyed items, forced-choice items, ranking items, items with various anchor labels so as to manipulate response scale usage and manipulation of item position. Additionally, varying item formats will be used within a questionnaire (e.g. within a single questionnaire, to mix 3- vs. 4- vs. 5-point Likert-scale items). Finally, the effects of using vignettes or situational judgment items, and also alternative framings for questions (e.g. "compared to other students in your class," vs. "in your country;" or "compared to other fields of study") were analysed. Other design procedures were also suggested, such as identifying a multi-lingual sample within a country that could be used to centre scale points. Some of these methods require the collection of new data, and therefore could only be accomplished during the field trial.

Extending the sample to study change at the school level (international option)

The need for longitudinal data at the school level becomes clear in the light of reverse causality (also called reciprocal determinism, simultaneous effects or recursive effects; Scheerens and Bosker, 1997). An example of reverse causality is when low performance causes certain policy decisions (e.g. remedial programmes). In such a situation, cross-sectional data can produce negative correlations between performance and policy measures, even when controlling for student background.

Scheerens and Bosker (1997) argued that longitudinal research at the school level could provide an empirical basis for examining reverse causality effects, but they did not find any such studies. Improvements in the methodology of school effectiveness research (e.g. Creemers and Kyriakides, 2008), and the availability of fine-graded assessment and evaluation data for schools (including trend data over several years) have greatly changed this field of educational research. Researchers are aware of the phenomenon, and it is understood that for this reason, cross-sectional surveys are of limited value in explaining why some schools have more desirable outcomes than other schools.

One way of dealing with this issue within PISA is to allow for longitudinal analyses at the school level, by sampling the same schools twice in different PISA cycles in order to study school-level changes (i.e. changes in teacher qualifications, student socio-economic composition, student attitudes and student performance – and how these are related over time).

Longitudinal extensions at the student level (recommended for future cycles)

Adding a longitudinal component to the PISA design setting would enable the examination of a number of interesting policy issues. Firstly it would allow for the estimation of value-added indicators, i.e. the effect of different variables on gains in student performance. While cross-sectional studies only examine relationships between variables, the direction of such effects can also be established with a longitudinal design. For example, the use of cross-lagged panel analysis or latent change models would allow for deciding whether self-efficacy has an effect on performance or whether performance has an effect on self-efficacy, or whether in fact reciprocal effects between both are found.

Some countries have already taken up the possibility of a longitudinal national extension of PISA, i.e. Australia (*Longitudinal Study of Australian Youth*), Canada (*Youth in Transition Survey*), the Czech Republic, Denmark, Germany, Switzerland (*Transitions from Education to Employment*) and Uruguay. Two approaches have been used:

- In Australia, Canada, Denmark, Switzerland, and Uruguay these studies have examined the transition from secondary to post-secondary education and/or to the labour market (e.g. Andersen, 2005; Bertschy et al., 2008; Looker and Thiessen, 2008; Zoido and Gluszynski, 2009). Results contribute to the validation of the PISA performance tests by showing that both reading and mathematical literacy results predict indicators of educational pathways (drop-out, graduation, post-secondary education, university, unemployment, income and so forth). This design further allows for the examination of effects of background and attitudinal variables on future life outcomes. For example, it has been shown that in Canada future educational pathways can be predicted by students' academic engagement, educational expectations, extra-curricular activities, peer educational support and parental support (Looker and Thiessen, 2008; Thiessen, 2007). In Australia, the individual perception of teacher-student relations and teacher morale in school (Curtis and McMillan, 2008) are associated with later student outcomes. If more countries took up this option, it would be possible to examine whether the effects found in Canada and Australia can be generalised to other countries, and whether system-level characteristics (e.g. second chance programmes) have an effect on students' educational paths as well.
- Another option is to examine the development of students' mathematics performance during secondary school, and the associations of school- and individual-level input and processes with latent change or growth in student competencies. Cross-sectional studies often find negative effects of different process variables, like homework support, remedial classes or teacher co-operation. In part, this may be explained by the fact that these measures are often used in a remedial way. In other words, especially those students with learning difficulties are offered extra learning time. Likewise, in schools with many disciplinary problems, teachers may cooperate more closely because they feel more in need of support. Using a longitudinal design would allow to disentangle remedial use of measures from the actual effects of programmes. More generally, it would improve the validity of conclusions drawn from analyses of the effects of policies and processes on student performance, and thus provide a more valid knowledge base for policy decisions.

Such a longitudinal design has already been realised by two German studies, the Project for the Analysis of Performance Development in Mathematics (PALMA) and PISA International Plus (PISA-I-Plus). These studies showed, for example, that the average student enjoyment in Grade 7 classrooms had a significant effect on teacher enjoyment in Grade 8, which, in turn, had an effect on student enjoyment in Grade 8, mediated by perceived teacher enthusiasm (Frenzel et al., 2009). The study also revealed that classroom management had a significant effect on mathematics competency in Grade 10, and that mathematics performance in Grade 10 was significantly higher in schools with many active and/or discipline-oriented teachers as compared with schools with many passive teachers (Prenzel et al., 2006).

Another example that has recently been discussed in the United States (Loveless, 2009) considers the issue of misplaced students (i.e. the negative effects of being enrolled in advanced courses on low-achieving students). If the critique put forward by Loveless (in this specific context) holds, there should be a negative interaction effect of enrolment in special physics courses and certain risk factors (e.g. migration status or low reading literacy) on mathematics performance. Within a cross-sectional design, this hypothesis cannot be tested properly. Longitudinal extensions would hence assist in answering complex policy questions.

SUMMARY

This framework for the context questionnaires in PISA 2012 has outlined how PISA can be developed further as a sustainable database for educational policy and research. To this end, the framework started with a review of the general purpose and policy relevance of PISA and outlined the general knowledge base stemming from research in educational effectiveness. As PISA 2012 has again mathematics as its major domain, consideration has also been given to the specific issues involved in the contexts for teaching and learning mathematics. However, the framework's centrepiece is its aim to map out a design for the PISA context questionnaires that will be sustainable well into the future. To this end, the framework puts a system in place that accommodates recurring general material that is covered in every cycle and domain-specific material, which is covered every fourth cycle, thus allowing for trend analyses of general as well as domain-specific issues. In addition, the framework's system also allows for thematic extensions and specific foci to enable PISA to anticipate and incorporate new material or topics of interest to its audience. Based on the analyses of PISA data and other research findings, the framework proceeds to make specific recommendations regarding material to be included in PISA 2012. The framework ends by making suggestions regarding aspects of design and analysis that are intended to build on the existing strengths of PISA and mapping an even stronger path into the future.

Notes

1. For more information, visit: *www.oecd.org/pisa/pisainfocus/*.

2. All these publications are available at: *http://www.oecd.org/pisa/pisaproducts/pisain-depthreports.htm.*

3. Abstracts and papers are available at *www.pisaresconf09.org/*.

4. In order to connect the present framework to earlier conceptual work on PISA, the sections "System-level inputs and processes", "School-level inputs and processes", "Classroom-level inputs and processes", and "Student-level inputs and processes" draw strongly on the *PISA 2009 Questionnaire Framework*, which was authored by Hank Levin and based on extensive work done by Jaap Scheerens. The authors gratefully acknowledge their important work as well as contributions made by other experts in previous PISA cycles.

5. Visit: *http://www.oecd.org/site/worldforum06/oecdindicatorinitiatives.htm.*

6. See EDU/PISA/GB (2005)21, p.17, par. 67.

7. Variables that were used in 2003 only, or dropped later, include family structure, educational career and reason for attending school.

8. All this information will be made available in the *PISA 2012 Technical Report (forthcoming)*.

References

Abedi J., M. Courtney, S. Leon, J. Kao and **T. Azzam** (2006), *English Language Learners and Math Achievement: A Study of Opportunity to Learn and Language Accommodation* (CSE Report 702, 2006), University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, California.

Ajzen, I. (1991), "The Theory of Planned Behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50, pp. 179-211.

Alexander, K.L., D.R. Entwisle and **L.S. Olson** (2007), "Lasting consequences of the summer learning gap", *American Sociological Review*, Vol. 72, pp. 167-180.

Amiot, C. and **R.Y. Bourhis** (2005), "Discrimination between dominant and subordinate groups: The positive-negative asymmetry effect and normative processes", *British Journal of Social Psychology*, Vol. 44, pp. 289-308.

Andersen, D. (2005), "Four Years after Basic Compulsory Education 19-year-olds on Post Compulsory Education Choices and Paths", *http://www.akf.dk/udgivelser_en/2005/4aarefter_grundskolen*.

Angelone, D., U. Moser and **E. Ramseier** (2009), "Instruction Time and Performance: Analyses of the Importance of Instruction Time for Competencies in Science and Mathematics based on Swiss PISA 2006 Data", paper presented at the PISA Research Conference, Kiel, Germany, 14-16 September, *www.pisaresconf09.org/user_uploads/files/context/room3/Angelone_Moser_Ramseier.pdf*.

Baker, D.P. (2009), "The Invisible Hand of World Education Culture", in G. Sykes, B. Schneider and D.N. Plank (eds.), *Handbook of Education Policy Research*, Routledge, New York, pp. 958-968.

Baker, D.P., B. Goesling and **G.K. LeTendre** (2002), "Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the 'Heyneman-Loxley effect' on Mathematics and Science Achievement", *Comparative Education Review*, Vol. 46, pp. 291-312.

Baker, D.P. and **G.K. LeTendre** (2005), *National Differences, Global Similarities: World Culture and the Future of Schooling*, Stanford University Press, Stanford, California.

Baker, M.L., J.N. Sigmon and **M.E. Nugent** (2001), "Truancy Reduction: Keeping Students in School", *Juvenile Justice Bulletin*, Office of Juvenile Justice and Delinquency Prevention, Washington D.C.

Bandura, A. (1977), *Social Learning Theory*, Prentice Hall, Englewood Cliffs, New Jersey.

Baumert J., C.H. Carstensen and **T. Siegle** (2005), ''Wirtschaftliche, Soziale und Kulturelle Lebensverhältnisse und Regionale Disparitäten des Kompetenzerwerbs", in PISA-Konsortium Deutschland (ed.), *PISA 2003: Der zweite Vergleich der Länder Deutschland – Was wissen und können Jugendliche?*, Waxmann Verlag, Münster.

Baumert, J., W. Blum, M. Brunner, T. Dubberke, A. Jordan, U. Klusmann, S. Krauss, M. Kunter, K. Löwen, M. Neubrand and **Y.-M. Tsai** (2008), *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz* (COACTIV) - *Dokumentation der Erhebungsinstrumente* (Baumert et al. (2009), Materialien aus der Bildungsforschung, No. 83), Max-Planck-Institut für Bildungsforschung, Berlin.

Baumert, J., M. Kunter, W. Blum, M. Brunner, T. Voss, A. Jordan, U. Klusmann, S. Krauss, M. Neubrand and **Y.-M. Tsai** (2010), "Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress", *American Educational Research Journal*, Vol. 47, pp. 133-180.

Baye, A., C. Monseur and **D. Lafontaine** (2009), "Institutional and Socioeconomic Factors Influencing the Reading Engagement", paper presented at the PISA Research Conference, Kiel, Germany, 14-16 September.

Belfield, C. and **H. Levin** (2002), "Education Privatization: Causes, Consequences, and Planning Implications", *International Institute of Educational Planning*, UNESCO, Paris.

Berezner, A. and **P. Lietz** (2009), "Implication for Preferred Model of Rotation on Cognitive Proficiency Estimate Generation Reporting and Subsequent Analyses", paper presented at the QEG Meeting, Offenbach, Germany, 19-21 October.

Berry, J.W. (1980), "Comparative Studies of Acculturative Stress", *International Migration Review*, Vol. 21, pp. 491-511.

Berry, J.W. (1990), "Psychology of Acculturation", in R.W. Brislin (ed.), *Applied Cross-Cultural Psychology*, Sage, Newbury Park, California, pp. 232-253.

Bertschy, K., E. Boni and **T. Meyer** (2008), "Young People in Transition from Education to Labor Market. Results of the Swiss youth panel survey TREE", update 2007, Basel.

Blase, J. and **J. Blase** (1998), *Handbook of Instructional Leadership: How Really Good Principals Promote Teaching and Learning*, Corwin Press, Thousand Oaks, California.

Blomhoj, M. and **T.H. Jensen** (2007), "What's all the Fuss about Competencies? Experiences with Using a Competence Perspective on Mathematics Education to Develop the Teaching of Mathematical Modelling", in W. Blum, P.L. Galbraith, H.-W. Henn and M. Niss (eds.), *Modelling and Applications in Mathematics Education*, 14 ICMI Study, Springer, New York, pp. 45-56.

Blum, W. and **D. Leiss** (2005), “How do Students and Teachers deal with Mathematical Modelling Problems? The Example ‘Sugarloaf’, in *ICTMA 12 Proceedings*, Ellis Horwood, Chichester.

Blum, W. and **D. Leiss** (2007), “Investigating Quality Mathematics Teaching – the DISUM Project”, in C. Bergsten and B. Grevholm (eds.), *Developing and Researching Quality in Mathematics Teaching and Learning. Proceedings of MADIF 5*, SMDF, Linköping, pp. 3-16.

Boaler, J. and **J. Greeno** (2000), “Identity, Agency and Knowing in Mathematical Worlds”, in J. Bowler (ed.), *Multiple perspectives on Mathematics Teaching and Learning*, Ablex Publishing, Westport, Connecticut.

Borko, H., K.L. Kuffner, S.C. Arnold, L. Creighton, B.M. Stecher, F. Martinez, D. Barnes and **M.L. Gilbert** (2007), *Using Artifacts to Describe Instruction: Lessons Learned from Studying Reform-Oriented Instruction in Middle School Mathematics and Science*, CSE Technical Report 705, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, California.

Bottani, N. and **B. Favre** (eds.) (2001), Open File: School Autonomy and Evaluation, *Prospects*, Vol. 31.

Britner, S.L. and **F. Pajares** (2006), “Sources of Science Self-Efficacy Beliefs of Middle School Students”, *Journal of Research in Science Teaching*, Vol. 43, pp. 485-499.

Bromme, R. (1997), “Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers“, in F. E. Weinert (ed.), Reihe: Enzyklopädie der Psychologie, *Pädagogische Psychologie*, Vol. 3, *Psychologie des Unterrichts und der Schule Hogrefe,* Göttingen, pp. 177-212.

Brookhart, S.M. (2009), *Exploring Formative Assessment*, ASCD Publishers, Alexandria, Virginia.

Brophy, J. (1992) (ed.), *Planning and Managing Learning Tasks and Activities: Advances in Research on Teaching*, Vol. 3, Greenwich, JAI Press, Connecticut.

Brown, A.L. (1994), “The Advancement of Learning”, *Educational Researcher*, Vol. 23, No. 8, pp. 4-12.

Brunner, M., M. Kunter, S. Krauss, U. Klusmann, J. Baumert, W. Blum, M. Neubrand, T. Dubberke, A. Jordan, K. Löwen and **Y.-M. Tsai** (2006), “Die professionelle Kompetenz von Mathematiklehrkräften: Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. Eine Zwischenbilanz des COACTIV-Projekts“, in M. Prenzel and L. Allolio-Näcke (eds.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*, Waxmann Verlag, Münster, pp. 54-82.

Buckley, J. (2009), *Cross-National Response Styles in International Educational Assessments: Evidence from PISA 2006*, Department of Humanities and Social Sciences in the Professions, Steinhardt School of Culture, Education, and Human Development, New York University, New York.

Carnoy, M., R. Elmore and **L. Siskin** (eds.) (2003), *The New Accountability: High Schools and High Stakes Testing,* Routledge Falmer, New York.

Carroll, J.B. (1963), “A Model of School Learning”, *Teachers College Record*, Vol. 64, pp. 723-733.

Christenson, S.L. (2004), “The Family-School Partnership: An Opportunity to Promote the Learning Competence of all Students”, *School Psychology Review*, Vol. 33, pp. 83-104.

Christiansen, B. and **G. Walther** (1986), “Task and Activity”, in B. Christiansen, A.G. Howson and M. Otte (eds.), *Perspectives on Mathematics Education*, pp. 243-307.

Chubb, J. and **T. Moe** (1990), *Politics, Markets, and American Schools*, The Brookings Institution, Washington, D.C.

Cleary, T.A. (1968), “Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges”, *Journal of Educational Measurement*, Vol. 5, pp. 115-124.

Coleman, J. (1988), “Social Capital in the Creation of Human Capital”, *American Journal of Sociology*, Vol. 94, pp. 95-120.

Common Core State Standards Initiative (CCSSI) (2010), Common Core State Standards, *http://www.corestandards.org/*.

Creemers, B.P.M. and **L. Kyriakides** (2008), *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice, and Theory in Contemporary Schools*, Routledge, London.

Curtis, D.D. and **J. McMillan** (2008), *School Non-completers: Profiles and Initial Destinations. Longitudinal Surveys of Australian Youth*, Research Report 54, The Australian Council for Educational Research Ltd., Camberwell, Victoria.

Desoete, A. and **M. Veenman** (eds.) (2006), *Metacognition in mathematics education*, Nova Science, Haupauge, New York.

Flavell, J.H. (1979), “Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry”, *American Psychologist*, Vol. 34, pp. 906-911.

Frenzel, A.C., T. Goetz, O. Lüdtke, R. Pekrun and **R.E. Sutton** (2009), “Emotional Transmission in the Classroom: Exploring the Relationship between Teacher and Student Enjoyment”, *Journal of Educational Psychology*, Vol. 101, pp. 705-716.

Fullan, M. (1992), *Successful School Improvement: The Implementation Perspective and Beyond*, Open University Press, Milton Keynes.

Fullan M. and **S. Stiegelbauer** (1991), *The New Meaning of Educational Change*, 2nd ed., Teachers College Press, New York.

Gamoran, A. (1992), "The Variable Effects of High School Tracking", *American Sociological Review*, Vol. 57, pp. 812-828.

Garofalo, J. and **F.K. Lester** (1985), "Metacognition, Cognitive Monitoring, and Mathematical Performance", *Journal for Research in Mathematics Education*, Vol. 16, pp. 163-176.

Gustafsson, J.E. (2007), "Understanding Causal Influences on Educational Achievement through Analysis of Differences over Time within Countries", in T. Loveless (ed.), *Lessons Learned: What International Assessments Tell Us about Math Achievement*, The Brookings Institution, Washington, D.C., pp. 37-63.

Hallinger, P. and **R.H. Heck** (1998), "Exploring the principal's contribution to school effectiveness: 1980 – 1995", *School Effectiveness and School Improvement*, Vol. 9, No. 2, pp. 157-191.

Hannaway, J. and **M. Carnoy** (eds.) (1993), *Decentralization and School Improvement: Can We Fulfill the Promise?*, Jossey-Bass, San Francisco, California.

Hanushek, E.A. and **S.G. Rivkin** (2007), "Teacher Quality", in E.A. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, pp. 1051-1078.

Hanushek, E.A. and **L. Woessmann** (2009), "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation", *NBER Working Paper No. 14633*, National Bureau of Economic Research, Cambridge, Massachusetts.

Hanushek, E.A. and **L. Woessmann** (2010), "The Economics of International Differences in Educational Achievement", *IZA Discussion Paper No. 4925*, IZA, Bonn.

Harris, A. and **J.H. Chrispeels** (eds.) (2006), *Improving Schools and Educational Systems: International Perspectives*, Routledge, London.

Heck, R.H. and **P. Hallinger** (2005), "The study of educational leadership and management. Where does the field stand today? ", *Educational Management Administration and Leadership*, Vol. 33, No. 2, pp. 229-244.

Hersbach, E. and **P. Lietz** (2010), "Towards a School Questionnaire in PISA 2012", QEG(1006)3.doc, paper presented at the QEG Meeting, Boston, Massachusetts, 21-23 June.

Hesse, H.G., K. Göbel and **J. Hartig** (2008), Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache", in DESI-Konsortium (eds.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch*, Weinheim, Beltz, pp. 208-230.

Hiebert J., R. Gallimore, H. Garnier, K.B. Givvin, H. Hollingsworth, J. Jacobs, A.M.-Y. Chui, D. Wearne, M. Smith, N. Kersting, A. Manaster, E. Tseng, W. Etterbeek, C. Manaster, P. Gonzales and **J. Stigler** (2003), *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study*, US Department of Education, National Center for Education Statistics, Washington, D.C.

Hill, H.C., B. Rowan and **D.L. Ball** (2005), "Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement", *American Educational Research Journal*, Vol. 42, pp. 371-406.

Hoover-Dempsey, K. and **H.M. Sandler** (1997), "Why do parents become involved in their children's education", *Review of Educational Research*, Vol. 67, pp. 3-42.

Hopkins, D. (ed.) (2005), *The Practice and Theory of School Improvement: International Handbook of Educational Change*, Springer, Dordrecht.

Hovey, J.D. and **C.A. King** (1996), "Acculturative stress, depression, and suicidal ideation among immigrant and second-generation Latino adolescents", *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 35, pp. 1183-1192.

Hui, C.H. and **H.C. Triandis** (1985), "Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies", *Journal of Cross-Cultural Psychology*, Vol. 16, No. 2, pp. 131-152.

Husén, T. (1967), *International Study of Achievement in Mathematics*, Vol. 2, Wiley, New York.

Jehanghir, K. and **C. Glas** (2007), "The Predictive Power of the Variables in the Student Questionnaires", paper presented at the QEG Meeting, Arnhem, 8-9 June.

Kao, G. and **M. Tienda** (1998), "Educational Aspirations of Minority Youth", *American Journal of Education*, Vol. 106, No. 3, pp. 349-384.

Kaplan, D. (2009a), "Advances in Multilevel Latent Variable Models for PISA Data", paper presented at the PISA Research Conference, Kiel, 14-16 September, *www.pisaresconf09.org/user_uploads/files/methodology/room1/Kaplan.pdf*.

Kaplan, D. (2009b), "Causal inference in non-experimental educational policy research", in G. Sykes, B. Schneider and D.N. Plank (eds.), Handbook on Education Policy Research, Taylor and Francis, New York, pp.139-153.

Kaplan, D. (2009c), *Structural Equation Modeling: Foundations and Extensions*, 2nd Edition, Sage Publications, Newbury Park, California.

Kaplan, D., J.-S. Kim and **S.-Y. Kim** (2009), "Multilevel Latent Variable Modeling: Current Research and Recent Developments", in R.E. Millsap (ed.), *The SAGE Handbook of Quantitative Methods in Psychology*, pp. 592-613.

Kearney, C.A. (2008), "School Absenteeism and School Refusal Behavior in Youth: A Contemporary Review", *Clinical Psychology Review*, Vol. 28, pp. 451-471.

Klieme, E., C. Pauli and **K. Reusser** (2009), "The Pythagoras Study: Investigating Effects of Teaching and Learning in Swiss and German Classrooms", in T. Janik and T. Seidel (eds.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, Waxmann Verlag, Münster, pp. 137-160.

Klieme, E., G. Schümer and **S. Knoll** (2001), "Mathematikunterricht in der Sekundarstufe I: Aufgabenkultur und Unterrichtsgestaltung", in Bundesministerium für Bildung und Forschung (BMBF) (ed.), *TIMSS-Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*, BMBF, Bonn, pp. 43-57.

Klieme, E. and **K. Rakoczy** (2003), "Unterrichtsqualität aus Schülerperspektive", in J. Baumert et al. (eds.), *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*, Leske und Budrich, Opladen, pp. 333-359.

Kobarg, M., T. Seidel, M. Prenzel, B. McCrae and **M. Walker** (2009), "Patterns of Science Teaching and Learning in an International Comparison", paper presented at the PISA Research Conference, Kiel, 14-16 September, *www.pisaresconf09.org/user_uploads/files/context/room3/Kobarg_etal.pdf*.

Koretz, D. (2008), *What Educational Testing Really Tells Us*, Harvard University Press, Cambridge, Massachusetts.

Krahn, H. and **A. Taylor** (2005), "Resilient teenagers: explaining the high educational aspirations of visible minority immigrant youth in Canada", *Journal of International Migration and Integration*, Vol. 6 (3/4), pp. 405-434.

Krüger, M.L., B. Witziers and **P.J.C. Sleegers** (2007), "The impact of school leadership on school level factors: Validation of a causal model", *School Effectiveness and School Improvement*, Vol. 18, No. 1, pp. 1-20.

Kunter, M. (2005), *Multiple Ziele im Mathematikunterricht*, Waxmann, Münster.

Kunter, M., Y.-M. Tsai, U. Klusmann, M. Brunner, S. Krauss and **J. Baumert (**2008), "Students' and Mathematics Teachers' Perception of Teacher Enthusiasm and Instruction", *Learning and Instruction*, Vol. 18, pp. 468-482.

Kyriakides, L. and **N. Tsangaridou** (2004), "School Effectiveness and Teacher Effectiveness in Physical Education", Paper presented at the *85 Annual AERA Meeting*, American Educational Research Association, Chicago, Illinois.

Lee, J. (2009), "Universals and Specifics of Math Self-Concept, Math Self-Efficacy, and Math Anxiety across 41 PISA 2003 Participating Countries", *Learning and Individual Differences*, Vol. 19, No. 3, pp. 355-365.

Lee, V.E. and **D.T. Burkam** (2003), "Dropping out of High-School: The Role of School-Organization and Structure", *American Educational Research Journal*, Vol. 40, No. 2, pp. 353-393.

Lee, V.E. and **J.B. Smith** (1997), "High School Size: Which Works Best and for Whom?", *Educational Evaluation and Policy Analysis*, Vol. 19, pp. 205-227.

Leithwood, K. and **C. Riehl** (2005), "What we know about successful school leadership", in W. Firestone and C. Riehl (eds.), *A new agenda: directions for research on educational leadership*, Teachers College Press, New York, pp.22-47.

Levin, H. (1978), "The Dilemma of Comprehensive Secondary School Reforms in Western Europe", *Comparative Education Review*, Vol. 22, pp. 434-451.

Libman, Z. (2009), "Teacher Licensing Examinations – True Progress or Illusion?", *Studies in Educational Evaluation*, Vol. 35, No. 1, pp. 7-15.

Lie, S. and **M. Kjaernsli** (2009), "Predictors for Students' Choice of Science Careers", paper presented at the PISA Research Conference, Kiel, Germany, 14-16 September, *www.pisaresconf09.org/user_uploads/files/context/room3/Lie_Kjaernsli.pdf.*

Liebkind, K. (1996), "Acculturation and stress: Vietnamese refugees in Finland", *Journal of Cross-Cultural Psychology*, Vol. 27, pp. 161-180.

Lipnevich, A.A., C. MacCann, S. Krumm, J. Burrus and **R.D. Roberts** (forthcoming), "Math Attitudes in Belarusian and US Middle School Students", *Journal of Educational Psychology*.

Lipowsky, F., K. Rakoczy, C. Pauli, B. Drollinger-Vetter, E. Klieme and **K. Reusser** (2009), "Quality of Geometry Instruction and its short-term Impact on Students' Understanding of the Pythagorean Theorem", *Learning and Instruction*, Vol. 19, pp. 527-537.

Looker, E.D. and **V. Thiessen** (2008), *The Second Chance System: Results from the three Cycles of the Youth in Transition Survey*, Learning Policy Directorate, Strategic Policy and Research, Quebec, *www.hrsdc.gc.ca/eng/publications_resources/learning_policy/sp-836-04-08e/sp_836_04_08e.pdf*.

Loveless, T. (2009), "How Well are American Students Learning? With Sections on International Assessments, the Misplaced Math Student, and Urban Schools", in *The 2008 Brown Center Report on American Education*, Brown Center on Education Policy at the Brookings Institute, Washington, D.C.

Lüdtke, O., H.W. Marsh, A. Robitzsch, U. Trautwein, T. Asparouhov and **B. Muthén** (2008), "The Multilevel Latent Covariate Model: A New, more Reliable Approach to Group-Level Effects in Contextual Studies", *Psychological Methods*, Vol. 13, pp. 203-229.

Ma, X. (1999), "Dropping out of advanced mathematics: The effects of parental involvement", *Teachers College Record*, Vol. 101, pp. 60-81.

Ma, X. and **N. Kishor** (1997), "Assessing the Relationship between Attitude toward Mathematics and Achievement in Mathematics: A Meta-Analysis", *Journal for Research in Mathematics Education*, Vol. 28, No. 1, pp. 26-47.

Mau, W.-C. and **L. Heim Bikos** (2000), "Educational and Vocational Aspirations of Minority and Female Students: A Longitudinal Study", *Journal of Counseling and Development*, Vol. 78, No. 2, pp. 186-194.

McCluskey, C.P., T.S. Bynum and **J.W. Patchin** (2004), "Reducing Chronic Absenteeism: An Assessment of an Early Truancy Initiative", *Crime and Delinquency*, Vol. 50, No. 2, p. 214-234.

McComb, E.M. and **C. Scott-Little** (2003), *A Review of Research on Participant Outcomes in After-School Programs: Implications for School Counselors*, ERIC Digest, ERIC Clearinghouse on Counseling and Student Services, ED482765, Greensboro, North Carolina.

Monk, D.H. (1994), "Subject Matter Preparation of Secondary Mathematics and Science Teachers and Student Achievement", *Economics of Education Review*, Vol. 13, No. 2, pp. 125-145.

Mullis, I.V.S. and **M.O. Martin** (2007), "TIMSS in Perspective: Lessons Learned from IEA's Four Decades of International Mathematics Assessments", in T. Loveless (ed.), *Lessons Learned. What International Assessments Tell Us about Math Achievement*, Brookings, Washington, pp. 9-36.

Muthén, B. (2002), "Beyond SEM: General Latent Variable Modeling", *Behaviormetrika*, Vol. 29, pp. 81-117.

National Council of Teachers for Mathematics (NCTM) (1989), *Curriculum and Evaluation Standards for School Mathematics*, NCTM, Reston, Virginia.

Niss, M. (2003), "Mathematical Competencies and the Learning of Mathematics: The Danish KOM Project", in A. Gagatsis and S. Papastavridis (eds.), *3rd Mediterranean Conference on Mathematical Education from 3-5 January 2003*, The Hellenic Mathematical Society, Athens, Hellas, pp. 115-124.

Oakes, J. (2005), *Keeping Track: How Schools Structure Inequality* (2nd ed.), Yale University, New Haven, Connecticut.

Opdenakker, M.-C. and **J. van Damme** (2000), "Effects of Schools, Teaching Staff and Classes on Achievement and Well-Being in Secondary Education: Similarities and Differences Between School Outcomes", *School Effectiveness and School Improvement*, Vol. 11, No. 2, pp. 165-196.

OECD (2001), *Knowledge and Skills for Life: First Results from PISA 2000*, PISA, OECD Publishing.

OECD (2004), *Learning for Tomorrow's World: First results from PISA 2003*, PISA, OECD Publishing.

OECD (2005a), *School Factors Related to Quality and Equity: Results from PISA 2000*, PISA, OECD Publishing.

OECD (2005b), *PISA 2003 Technical Report*, PISA, OECD Publishing.

OECD (2007a), *Education at a Glance: OECD Indicators*, OECD Publishing.

OECD (2007b), *PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis*, PISA, OECD Publishing

OECD (2008), *Education at a Glance: OECD Indicators*, OECD Publishing.

OECD (2009a), *Education at a Glance: OECD Indicators*, OECD Publishing.

OECD (2009b), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD Publishing.

OECD (2010), *PISA 2009 Results: Volumes I, II, II, IV, V and VI*, PISA, OECD Publishing.

OECD (2011), *Education at a Glance: OECD Indicators*, OECD Publishing.

OECD (2012a), *Education at a Glance: OECD Indicators*, OECD Publishing

OECD (2012b), *Untapped Skills: Realising the Potential of Immigrant Students*, PISA, OECD Publishing.

OECD (2012c), *Public and Private Schools: How management and funding relate to their socio-economic profile*, PISA, OECD Publishing.

OECD (2012d), *Let's read them a story! The parent factor in education*, PISA, OECD Publishing.

Pajares, F. and **M.D. Miller** (1997), Mathematics Self-Efficacy and Mathematical Problem Solving: Implications of Using Different Forms of Assessment, *The Journal of Experimental Education*, Vol. 65, pp. 213-228.

Perreault, S. and **R.Y. Bourhis** (1999), "Ethnocentrism, social identification and discrimination", *Personality and Social Psychology Bulletin*, Vol. 25, pp. 92-103.

Prenzel, M., J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost and **U. Schiefele** (eds.) (2006), *PISA 2003 : Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*, Waxmann Verlag, Münster.

Purves, A.C. (1987), "The Evolution of the IEA: A Memoir", *Comparative Education Review*, Vol. 31, No. 1, pp. 10-28.

Rakoczy, K., A. Buff and **F. Lipowsky** (2005), "Teil 1: Befragungsinstrumente", in E. Klieme et al. (eds.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie Unterrichtsqualität, Lernverhalten und mathematisches Verständnis*, GFPF, Frankfurt.

Rakoczy, K., E. Klieme and **C. Pauli** (2008), Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation [The impact of the perceived support of three basic psychological needs and of the perceived relevance of contents for students' self-determined motivation in mathematics instruction], *Zeitschrift für Pädagogische Psychologie*, Vol. 22, No. 1, pp. 25-35.

Raudenbush, S.W. and **A.S. Bryk** (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, Thousand Oaks, California.

Rost, J., C.H. Carstensen and **M. von Davier** (1997), "Applying the Mixed Rasch Model to Personality Questionnaires", in J. Rost and R. Langeheine (eds.), *Applications of latent trait and latent class models in the social sciences*, Waxmann, New York, pp. 324-332.

Rowan, B., E. Camburn and **R. Correnti** (2004), "Using teacher logs to measure the enacted curriculum in large-scale surveys: Insights from the Study of Instructional Improvement", *Elementary School Journal*, Vol. 105, pp. 75-102.

Ryan, R.M. and **E.L. Deci** (2000), Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions, *Contemporary Educational Psychology*, Vol. 25, pp. 54-67.

Rychen, D.S. and **L.H. Salganik** (eds.) (2003), *Key Competencies for a Successful Life and a Well-Functioning Society*, Hogrefe and Huber Publishers, Göttingen.

Scheerens, J. (2000), "Improving School Effectiveness", *Fundamentals of Educational Planning Series*, IIEP, Vol. 68, UNESCO, Paris.

Scheerens, J. and **R.J. Bosker** (1997), *The Foundations of Educational Effectiveness*, Pergamon, Oxford.

Schmid, C.L. (2001), "Educational Achievement, Language-Minority Students, and the New Second Generation", *Sociology of Education (Extra Issue)*, pp. 71-87.

Schmidt, W.H. and **C.C. McKnight** (1998), "What Can We Really Learn from TIMSS?", *Science*, Vol. 282, pp. 1831-1839.

Schmidt, W.H., C.C. McKnight, R.T. Houang, H.C. Wang, D.E. Wiley, L.S. Cogan and **R.G. Wolfe** (2001), *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*, Jossey Bass, San Francisco, California.

Schmidt, W.H. and **R.T. Houang** (2007), "Lack of Focus in the Mathematics Curriculum: Symptom or Cause", in T. Loveless (ed.), *Lessons Learned: What International Assessments Tell Us about Math Achievement*, Brookings, Washington, pp. 65-84.

Schmidt, W.H. and **A. Maier** (2009), "Opportunity to Learn", in G. Sykes, B. Schneider and D.N. Plank (eds.), *Handbook of Education Policy Research*, Routledge, New York, pp. 541-559.

Schmidt, W.H. and **C. McKnight** (1995), "Surveying Educational Opportunity in Mathematics and Science: An International Perspective", *Educational Evaluation and Policy Analysis*, Vol. 17, No. 3, pp. 337-353.

Schneider, W. and **C. Artelt** (forthcoming), "Metacognition and Mathematics Education", in *ZDM/ The International Journal on Mathematics Education*.

Schoenfeld, A.H. (ed.) (1987), *Cognitive Science and Mathematics Education*, Erlbaum, Hillsdale, New Jersey.

Schoenfeld, A.H. (1992), "Learning to Think Mathematically: Problem Solving, Metacognition, and Sense-Making in Mathematics", in D. Grouws, (ed.), *Handbook for Research on Mathematics Teaching and Learning*, MacMillan, New York, pp. 334-370.

Schoenfeld, A.H. (2006), "Mathematics Teaching and Learning", in P.A. Alexander and I.H. Winne (eds.), *Second Handbook of Educational Psychology*, Routledge, London.

Shulman, L. (1987), "Knowledge and Teaching: Foundations of the New Reform", *in Harvard Educational Review*, Vol. 57, No. 1, pp. 1-22.

Slavin, R.E. (1990), "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis", in *Review of Educational Research*, Vol. 60, No. 3, pp. 471-499.

Snijders, T. and **R. Bosker,** (1999), *Multilevel Analysis*, Sage Publications, London.

Stanat, P. (2006), "Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft", in J. Baumert, P. Stanat and R. Watermann (ed.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit; vertiefende Analysen im Rahmen von PISA 2000*, VS Verl. für Sozialwiss, Wiesbaden, pp. 189-219.

Stanat, P. and **G. Christensen,** (2006), *Where Immigrant Students Succeed - A Comparative Review of Performance and Engagement in PISA 2003*, OECD Publishing.

Stein, M.K., R.A. Engle, M.S. Smith and **E.K. Hughes** (2008), "Orchestrating Productive Mathematical Discussions: Five Practices for helping Teachers move beyond show and tell", in *Mathematical Thinking and Learning*, Vol. 10, pp. 313-340.

Stevens, F. (1993), "Applying an Opportunity-to-Learn Conceptual Framework to the Investigation of the Effects of Teaching Practices via Secondary Analyses of Multiple-Case-Study Summary Data", in *Journal of Negro Education*, Vol. 62, No. 3, pp. 232-248.

Stigler, J.W. and **J. Hiebert** (1999), *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*, Free Press, New York.

Sui-Chu, H. and **J.D. Willms** (1996), "Effects of parental involvement on eighth-grade achievement", *Sociology of Education*, Vol. 69, pp. 126-141.

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

Thiessen, V. (2007), *The Impact of Factors on Trajectories that Lead to a High School Diploma and to Participation in Post Secondary Education*, Human Resources and Social Development Canada Publications Centre, Quebec.

Turner, S.L., J.C. Steward and **R.T. Lapan (**2004), "Family Factors Associated with Sixth Grade Adolescents' Math and Science Career Interests", *Career Development Quarterly*, Vol. 53, No. 1, pp. 41-52.

Van de Gaer, E. and **R. Adams** (2010), "The Modeling of Response Style Bias: An Answer to the Attitude-Achievement Paradox?", paper presented at the annual conference of the American Educational Research Association, Denver, Colorado, 30 April-4 May.

Van de Vijver, F.J.R., D.A. Van Hemert and **Y.H. Poortinga** (eds.) (2008), *Multilevel Analysis of Individuals and Cultures,* Erlbaum, New York.

Van Landeghem, G., B. De Fraine and **J. Van Damme** (2005), "The Consequence of Ignoring a Level of Nesting in Multilevel Analysis: A Comment", *Multivariate Behavioral Research*, Vol. 40, pp. 423-434.

Vieluf, S., J. Lee and **P. Kyllonen** (2009a), "The Predictive Power of Variables from the PISA 2003 Student Questionnaire", paper presented at the QEG Meeting, Offenbach, Germany, 19-21 October.

Vieluf, S., J. Lee and **P. Kyllonen** (2009b), "The Cross-Cultural Validity of Variables from the PISA 2003 Student Questionnaire", paper presented at the QEG Meeting, Offenbach, Germany, 19-21 October.

Walberg, H.J. (1986), "Syntheses of Research on Teaching", in M. C. Wittrock (ed.), in *Handbook of research on teaching* (3rd ed.), Macmillan, New York, pp. 214-229.

Wang, M.C., G.D. Haertel and **H.D. Walberg** (1993), "Toward a Knowledge Base for School Learning", *in Review of Educational Research*, Vol. 63, No. 3, pp. 249 -294.

Wayne, A. and **P. Youngs** (2003), "Teacher Characteristics and Student Achievement Gain: A Review", *in Review of Educational Research*, Vol. 73, pp. 89-122.

Wigfield, A., J.S. Eccles and **D. Rodriguez** (1998), "The Development of Children's Motivation in School Contexts", in A. Iran-Nejad and P.D. Pearson (eds.), in *Review of Research in Education*, American Educational Research Association, Vol. 23, Washington D.C.

Wilmers, N., D. Enzmann, D. Schaefer, K. Herbers, W. Greve and **P. Wetzels** (2002), *Jugendliche in Deutschland zur Jahrtausendwende: Gefährlich oder gefährdet? Ergebnisse wiederholter, repräsentativer Dunkelfelduntersuchungen zu Gewalt und Kriminalität im Leben junger Menschen 1998 – 2000*, Nomos Verlagsgesellschaft, Baden-Baden, Germany.

Witziers, B., R.J. Bosker and **M.L. Krüger** (2003), "Educational leadership and student achievement: The elusive search for an association", *Educational Administrative Quarterly*, Vol. 39, No. 3, pp. 398-425.

Woessmann, L. (2006), "The Complementarity of Central Exams and School Autonomy: Economic Theory and International Evidence", in E. Gori, D. Vidoni, E. Hanushek and C. Glenn (eds.), *Institutional Models in Education: Legal Framework and Methodological Aspects for a New Approach to the Problem of School Governance*, Wolf Legal Publishers Nijmegen, pp. 47-65.

Yasushi, O. (2009), "Comparison of Attitudes toward Science between Grade 9 and 10 Japanese Students by Using the PISA Questions and Its Implications on Science Teaching in Japan", paper presented at the QEG Meeting, Offenbach, Germany, 19-21 October.

Zoido, P. and **T. Gluszynski** (2009), "Pathways to Success: How Knowledge and Skills at Age 15 Shape Future Life Outcomes", paper presented at the PISA Research Conference, Kiel, Germany, 14-16 September.

Zuzovsky, R. (2009), "Teachers' Qualifications and Their Impact on Student Achievement: Findings from TIMSS 2003 Data for Israel", in M. von Davier and D. Hastedt (eds.), *IERI Monograph Series Issues and Methodologies in Large-Scale Assessments*, IER Institute, Hamburg / Princeton, Germany / USA, Vol. 2, pp. 37-62.

Annex A

BACKGROUND QUESTIONNAIRES

Annex A provides the background questionnaires used in the PISA 2012 survey to obtain information about the participating schools, students and parents.

The school questionnaire is administered to school principals and covers the structure and organisation of: the school; the student and teacher body; the school's resources; the school's instruction, curriculum and assessment; the school climate; the school's policies and practices; financial education at school, and comprises one additional question for the online school questionnaire.

The student questionnaire addresses: the student characteristics and educational career; family context and home resources; mathematics learning; experience with different kinds of mathematics problems at school; mathematics experiences; classroom and school climate; and problem solving experiences.

Two optional questionnaires are administered to students: an educational career questionnaire covering the students' educational histories, their career aspirations and questions on language learning; and a questionnaire about students' access to information and communication technologies (ICTs) and student's use of ICTs and attitude towards computers.

A third optional questionnaire for parents covers: parental background; the cost of educational services; parents' attitudes to their child's school; parents' involvement with school; school choice; parental support for learning at home; mathematics in child's future career and job market; academic and professional expectations in mathematics; child's past academic performance; child's career interests; and parents' immigration background.

SCHOOL QUESTIONNAIRE

The school questionnaire is administered to the school principal and takes about 30 minutes to complete. It covers a variety of school-related aspects:

- Structure and organisation of the school
- Student and teacher body
- School's resources
- School's instruction, curriculum and assessment
- School climate
- School's policies and practices
- Financial education at school
- Additional question for ONLINE school questionnaire

Technical terms are given in <brackets> and are adapted to the national context by the national data collection centre of the participating country or economy. In this annex, an explanation of the technical terms is given below the questionnaire item.

The structure and organisation of the school

Q1 (SC01) **Is your school a public or a private school?**
(Please tick only one box.)

A public school *(This is a school managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise.)*	☐1
A private school *(This is a school managed directly or indirectly by a non-government organisation; e.g. a church, trade union, business, or other private institution.)*	☐2

Q2 (SC02) **About what percentage of your total funding for a typical school year comes from the following sources?**
(Please write a number in each row. Write 0 (zero) if no funding comes from that source.)

		%
a)	Government (includes departments, local, regional, state and national)	
b)	Student fees or school charges paid by parents	
c)	Benefactors, donations, bequests, sponsorships, parent fundraising	
d)	Other	
	Total	**100 %**

Q3 (SC03) **Which of the following definitions best describes the community in which your school is located?**
(Please tick only one box.)

A village, hamlet or rural area (fewer than 3 000 people)	☐1
A small town (3 000 to about 15 000 people)	☐2
A town (15 000 to about 100 000 people)	☐3
A city (100 000 to about 1 000 000 people)	☐4
A large city (with over 1 000 000 people)	☐5

Q4 (SC04)

We are interested in the options parents have when choosing a school for their children.

Which of the following statements best describes the schooling available to students in your location?

(Please tick only one box.)

There are two or more other schools in this area that compete for our students	☐ 1
There is one other school in this area that competes for our students	☐ 2
There are no other schools in this area that compete for our students	☐ 3

Q5 (SC05)

What is the average size of <test language> classes in <national modal grade for 15-year-olds> in your school?

(Please tick only one box.)

15 students or fewer	☐ 1
16-20 students	☐ 2
21-25 students	☐ 3
26-30 students	☐ 4
31-35 students	☐ 5
36-40 students	☐ 6
41-45 students	☐ 7
46-50 students	☐ 8
More than 50 students	☐ 9

Notes: <**national modal grade for 15-year-olds**> is the name of the grade attended by most 15-year-olds in the participating country or economy.

<**test language**> is replaced with the name of the language used in the PISA mathematics literacy test.

The student and teacher body

Q6 (SC07)

As at <month day, 2012>, what was the total school enrolment (number of students)?

(Please write a number on each line. Write 0 (zero) if there are none.)

a)	Number of boys	
b)	Number of girls	

Note: <**month day, 2012**> should be a date about one month before the data collection.

Q7 (SC09)

How many of the following teachers are on the staff of your school?

Include both full-time and part-time teachers*. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.*

(Please write a number in each space provided. Write 0 (zero) if there are none.)

		Full-time	Part-time
a)	Teachers in TOTAL		
b)	Teachers fully certified by <the appropriate authority>		
c)	Teachers with an <ISCED 5A> qualification		

Notes: <**the appropriate authority**> refers to the government agency which is empowered to certify that a person is permitted to work as a school teacher.

<**ISCED 5A**> refers to qualification obtained from a tertiary study programme with a strong theoretical foundation typically with a minimum duration of three years' full time equivalent, providing entry into a profession with high skills requirements or an advanced research programme.

Q8 (SC10)

How many of the following are on the <mathematics staff> of your school?

Include both full-time and part-time teachers*. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.*

Please count only those teachers who have taught or will teach mathematics during the current school year.

(Please write a number in each space provided. Write 0 (zero) if there are none.)

		Full-time	Part-time
a)	Teachers of mathematics in TOTAL		
b)	Teachers of mathematics with an <ISCED 5A> qualification		
c)	Teachers of mathematics with an <ISCED 5A> qualification <with a major> in mathematics		
d)	Teachers of mathematics with an <ISCED 5A> qualification in <pedagogy>		
e)	Teachers of mathematics with an <ISCED 5B> but not an <ISCED 5A> qualification		

Notes: For a definition of <**ISCED 5A**> see Q7.

<**ISCED 5B**> refers to qualification obtained in tertiary programmes that are generally more practical/technical/occupationally specific and typically shorter than ISCED 5A programmes. Typically, these programmes have a minimum of two years' full-time equivalent duration and prepare students to enter a particular occupation.

<**with a major**> refers to the focus of study in an undergraduate university degree. A major in mathematics is a complete sequence of mathematics in an ISCED 5A qualification.

The school's resources

The goal of the following set of three questions is to gather information about the student-computer ratio for students in the <national modal grade for 15-year-olds> at your school.

		Number
Q9a (SC11)	**At your school, what is the total number of students in the <national modal grade for 15-year-olds>?**	
Q9b (SC11)	**Approximately, how many computers are available for these students for educational purposes?**	
Q9c (SC11)	**Approximately, how many of these computers are connected to the Internet/ World Wide Web?**	

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q10 (SC13)

In all subjects taken together, for how much of the work does the school expect <national modal grade for 15-year-olds> students to access the Internet/ World Wide Web?

(Please tick only one box in each row.)

		<10%	10-25%	26-50%	51-75%	>75%
a)	Work during lessons	☐1	☐2	☐3	☐4	☐5
b)	Homework	☐1	☐2	☐3	☐4	☐5
c)	Assignments or projects	☐1	☐2	☐3	☐4	☐5

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q11 (SC14) **Is your school's capacity to provide instruction hindered by any of the following issues?**
(Please tick one box in each row.)

		Not at all	Very little	To some extent	A lot
a)	A lack of qualified science teachers	☐1	☐2	☐3	☐4
b)	A lack of qualified mathematics teachers	☐1	☐2	☐3	☐4
c)	A lack of qualified <test language> teachers	☐1	☐2	☐3	☐4
d)	A lack of qualified teachers of other subjects	☐1	☐2	☐3	☐4
e)	Shortage or inadequacy of science laboratory equipment	☐1	☐2	☐3	☐4
f)	Shortage or inadequacy of instructional materials (e.g. textbooks)	☐1	☐2	☐3	☐4
g)	Shortage or inadequacy of computers for instruction	☐1	☐2	☐3	☐4
h)	Lack or inadequacy of Internet connectivity	☐1	☐2	☐3	☐4
i)	Shortage or inadequacy of computer software for instruction	☐1	☐2	☐3	☐4
j)	Shortage or inadequacy of library materials	☐1	☐2	☐3	☐4
k)	Shortage or inadequacy of school buildings and grounds	☐1	☐2	☐3	☐4
l)	Shortage or inadequacy of heating/cooling and lighting systems	☐1	☐2	☐3	☐4
m)	Shortage or inadequacy of instructional space (e.g. classrooms)	☐1	☐2	☐3	☐4

Note: For a definition of <**test language**> see Q5.

School instruction curriculum and assessment

Q12 (SC15) *Schools sometimes organise instruction differently for students with different abilities and interests in mathematics.*
Which of the following options describe what your school does for <national modal grade for 15-year-olds> students in mathematics classes?
(Please tick one box in each row.)

		For all classes	For some classes	Not for any classes
a)	Mathematics classes study similar content, but at different levels of difficulty	☐1	☐2	☐3
b)	Different classes study different content or sets of mathematics topics that have different levels of difficulty	☐1	☐2	☐3
c)	Students are grouped by ability within their mathematics classes	☐1	☐2	☐3
d)	In mathematics classes, teachers use pedagogy suitable for students with heterogeneous abilities (i.e. students are not grouped by ability)	☐1	☐2	☐3

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q13 (SC16) **<This academic year>, which of the following activities does your school offer to students in the <national modal grade for 15-years-olds>?**
(Please tick one box in each row.)

		Yes	No
a)	Band, orchestra or choir	☐1	☐2
b)	School play or school musical	☐1	☐2
c)	School yearbook, newspaper or magazine	☐1	☐2
d)	Volunteering or service activities, e.g. <national examples>	☐1	☐2
e)	Mathematics club	☐1	☐2
f)	Mathematics competitions, e.g. <national examples>	☐1	☐2
g)	Chess club	☐1	☐2
h)	Club with a focus on computers/ Information and Communication Technology	☐1	☐2
i)	Art club or art activities	☐1	☐2
j)	Sporting team or sporting activities	☐1	☐2
k)	<country specific item>	☐1	☐2

Notes: <**This academic year**> refers to the year of schooling which is not necessarily the calendar year.
For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q14 (SC18) **In your school, are assessments of students in <national modal grade for 15-year-olds> used for any of the following purposes?**
(Please tick one box in each row.)

		Yes	No
a)	To inform parents about their child's progress	☐1	☐2
b)	To make decisions about students' retention or promotion	☐1	☐2
c)	To group students for instructional purposes	☐1	☐2
d)	To compare the school to <district or national> performance	☐1	☐2
e)	To monitor the school's progress from year to year	☐1	☐2
f)	To make judgements about teachers' effectiveness	☐1	☐2
g)	To identify aspects of instruction or the curriculum that could be improved	☐1	☐2
h)	To compare the school with other schools	☐1	☐2

Notes: For a definition of <**national modal grade for 15-year-olds**> see Q5.
<**district or national**> performance refers to comparison with a larger administrative region which could be the district, region, province and/or the country as a whole.

Q15 (SC19) **In your school, are achievement data used in any of the following <accountability procedures>?**
*Achievement data include **aggregated** school or grade-level test scores or grades, or graduation rates.*
(Please tick one box in each row.)

		Yes	No
a)	Achievement data are posted publicly (e.g. in the media)	☐1	☐2
b)	Achievement data are tracked over time by an administrative authority	☐1	☐2

Note: <**accountability procedures**> means the regular use of school-level statistics on student achievement to report on the quality of the school functioning to parents or external authorities.

Q16 (SC20) **Does your school offer mathematics lessons in addition to the mathematics lessons offered during the usual school hours?**

(Please tick only one box.)

Yes	☐1	*go to the next question*
No	☐2	*go to Q18*

Q17 (SC21) **What is the purpose of these additional mathematics lessons?**

(Please tick only one box.)

<Enrichment mathematics> only	☐1
<Remedial mathematics> only	☐2
Both <enrichment mathematics> and <remedial mathematics>	☐3
Without differentiation depending on the prior achievement level of the students	☐4

Notes: <**Enrichment mathematics**> is mathematics offered <u>outside of normal class time</u> to extend/stimulate/challenge students who are of higher ability. <**Remedial mathematics**> is mathematics offered <u>outside of normal class time</u> to help students who have fallen behind the performance level of their peers to catch up.

School climate

Q18 (SC22) **In your school, to what extent is the learning of students hindered by the following phenomena?**

(Please tick one box in each row.)

		Not at all	Very little	To some extent	A lot
a)	Student truancy	☐1	☐2	☐3	☐4
b)	Students skipping classes	☐1	☐2	☐3	☐4
c)	Students arriving late for school	☐1	☐2	☐3	☐4
d)	Students not attending compulsory school events (e.g. sports day) or excursions	☐1	☐2	☐3	☐4
e)	Students lacking respect for teachers	☐1	☐2	☐3	☐4
f)	Disruption of classes by students	☐1	☐2	☐3	☐4
g)	Student use of alcohol or illegal drugs	☐1	☐2	☐3	☐4
h)	Students intimidating or bullying other students	☐1	☐2	☐3	☐4
i)	Students not being encouraged to achieve their full potential	☐1	☐2	☐3	☐4
j)	Poor student-teacher relations	☐1	☐2	☐3	☐4
k)	Teachers having to teach students of heterogeneous ability levels within the same class	☐1	☐2	☐3	☐4
l)	Teachers having to teach students of diverse ethnic backgrounds (i.e. language, culture) within the same class	☐1	☐2	☐3	☐4
m)	Teachers' low expectations of students	☐1	☐2	☐3	☐4
n)	Teachers not meeting individual students' needs	☐1	☐2	☐3	☐4
o)	Teacher absenteeism	☐1	☐2	☐3	☐4
p)	Staff resisting change	☐1	☐2	☐3	☐4
q)	Teachers being too strict with students	☐1	☐2	☐3	☐4
r)	Teachers being late for classes	☐1	☐2	☐3	☐4
s)	Teachers not being well prepared for classes	☐1	☐2	☐3	☐4

Note: **'Student truancy' is used differently** in PISA 2012 and refers ONLY TO THE UNAUTHORISED failure to attend classes, whereas in previous cycles 'student absenteeism' included the unauthorised AND authorised (e.g. illness) absence of students from school.

Q19 (SC23)	During <the last academic year>, what proportion of students left your school without a <certificate or qualification that allows students to enter post-school destinations such as university, technical, further or vocational education, apprenticeships or employment>?	%

Note: <**the last academic year**> refers to the previous year of schooling, not necessarily the previous calendar year.

Q20 (SC24) **Which statement below best characterises parental expectations towards your school?**
(Please tick only one box.)

	There is *constant pressure* from many parents, who expect our school to set very high academic standards and to have our students achieve them	☐1
	Pressure on the school to achieve higher academic standards among students comes from a *minority of parents*	☐2
	Pressure from parents on the school to achieve higher academic standards among students is *largely absent*	☐3

Q21 (SC25) **During <the last academic year>, what proportion of students' parents participated in the following school-related activities?**
(Please write a number in each row. Write 0 (zero) if no parents participated in the activity. Write 100 (one hundred) if all parents participated in the activity.)

		%
a)	Discussed their child's behaviour with a teacher on their own initiative	
b)	Discussed their child's behaviour on the initiative of one of their child's teachers	
c)	Discussed their child's progress with a teacher on their own initiative	
d)	Discussed their child's progress on the initiative of one of their child's teachers	
e)	Volunteered in physical activities, e.g. building maintenance, carpentry, gardening or yard work	
f)	Volunteered in extra-curricular activities, e.g. book club, school play, sports, field trip	
g)	Volunteered in the school library or media centre	
h)	Assisted a teacher in the school	
i)	Appeared as a guest speaker	
j)	Participated in local school <government>, e.g. parent council or school management committee	
k)	Assisted in fundraising for the school	
l)	Volunteered in the school <canteen>	

Note: For a definition of <**the last academic year**> see Q19.

Q22 (SC26) **Think about the teachers in your school. How much do you agree with the following statements?**
(Please tick one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	The morale of teachers in this school is high	☐1	☐2	☐3	☐4
b)	Teachers work with enthusiasm	☐1	☐2	☐3	☐4
c)	Teachers take pride in this school	☐1	☐2	☐3	☐4
d)	Teachers value academic achievement	☐1	☐2	☐3	☐4

Q23 (SC27) **How much do you agree with these statements about teachers in your school?**
(Please tick one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Mathematics teachers are interested in trying new methods and teaching practices	☐1	☐2	☐3	☐4
b)	There is a preference among mathematics teachers to stay with well-known methods and practices	☐1	☐2	☐3	☐4

Q24 (SC28) **How much do you agree with these statements about teachers in your school?**
(Please tick one box in each row.)

c)	There is consensus among mathematics teachers that academic achievement must be kept as high as possible	☐1	☐2	☐3	☐4
d)	There is consensus among mathematics teachers that it is best to adapt academic standards to the students' levels and needs	☐1	☐2	☐3	☐4

Q25 (SC29) **How much do you agree with these statements about teachers in your school?**
(Please tick one box in each row.)

e)	There is consensus among mathematics teachers that the social and emotional development of the students is as important as their acquisition of mathematical skills and knowledge in mathematics classes	☐1	☐2	☐3	☐4
f)	There is consensus among mathematics teachers that the development of mathematical skills and knowledge in students is the most important objective in mathematics classes	☐1	☐2	☐3	☐4

Q26 (SC30) **During the last year, have any of the following methods been used to monitor the practice of mathematics teachers at your school?**
(Please tick one box in each row.)

		Yes	No
a)	Tests or assessments of student achievement	☐1	☐2
b)	Teacher peer review (of lesson plans, assessment instruments, lessons)	☐1	☐2
c)	Principal or senior staff observations of lessons	☐1	☐2
d)	Observation of classes by inspectors or other persons external to the school	☐1	☐2

Q27 (SC31) **To what extent have appraisals of and/or feedback to teachers directly led to the following?**
(Please tick one box in each row.)

		No change	A small change	A moderate change	A large change
a)	A change in salary	☐1	☐2	☐3	☐4
b)	A financial bonus or another kind of monetary reward	☐1	☐2	☐3	☐4
c)	Opportunities for professional development activities	☐1	☐2	☐3	☐4
d)	A change in the likelihood of career advancement	☐1	☐2	☐3	☐4
e)	Public recognition from you	☐1	☐2	☐3	☐4
f)	Changes in work responsibilities that make the job more attractive	☐1	☐2	☐3	☐4
g)	A role in school development initiatives (e.g. curriculum development group, development of school objectives)	☐1	☐2	☐3	☐4

School policies and practices

Q28 (SC32) **How often are the following factors considered when students are admitted to your school?**
(Please tick one box in each row.)

		Never	Sometimes	Always
a)	Student's record of academic performance (including placement tests)	☐1	☐2	☐3
b)	Recommendation of feeder schools	☐1	☐2	☐3
c)	Parents' endorsement of the instructional or religious philosophy of the school	☐1	☐2	☐3
d)	Whether the student requires or is interested in a special programme	☐1	☐2	☐3
e)	Preference given to family members of current or former students	☐1	☐2	☐3
f)	Residence in a particular area	☐1	☐2	☐3
g)	Other	☐1	☐2	☐3

Q29 (SC33) **Regarding your school, who has a considerable responsibility for the following tasks?**
(Please tick as many boxes as appropriate in each row.)

		Principal	Teachers	<School governing board>	<Regional or local education authority>	National education authority
a)	Selecting teachers for hire	☐1	☐2	☐3	☐4	☐5
b)	Firing teachers	☐1	☐2	☐3	☐4	☐5
c)	Establishing teachers' starting salaries	☐1	☐2	☐3	☐4	☐5
d)	Determining teachers' salary increases	☐1	☐2	☐3	☐4	☐5
e)	Formulating the school budget	☐1	☐2	☐3	☐4	☐5
f)	Deciding on budget allocations within the school	☐1	☐2	☐3	☐4	☐5
g)	Establishing student disciplinary policies	☐1	☐2	☐3	☐4	☐5
h)	Establishing student assessment policies	☐1	☐2	☐3	☐4	☐5
i)	Approving students for admission to the school	☐1	☐2	☐3	☐4	☐5
j)	Choosing which textbooks are used	☐1	☐2	☐3	☐4	☐5
k)	Determining course content	☐1	☐2	☐3	☐4	☐5
l)	Deciding which courses are offered	☐1	☐2	☐3	☐4	☐5

Notes: <**school governing board**> is a board directly responsible for the governance of the school. This board may be totally external to the school or may have staff and student representation.
<**Regional or local education authority**> is an authority that is not a national authority and does not directly govern the school.

Q30 (SC34) **Below are statements about your management of this school. Please indicate the frequency of the following activities and behaviours in your school during <the last academic year>.**
(Please tick only one box in each row.)

		Did not occur	1-2 times during the year	3-4 times during the year	Once a month	Once a week	More than once a week
a)	I work to enhance the school's reputation in the community	☐1	☐2	☐3	☐4	☐5	☐6
b)	I use student performance results to develop the school's educational goals	☐1	☐2	☐3	☐4	☐5	☐6
c)	I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school	☐1	☐2	☐3	☐4	☐5	☐6
d)	I ensure that teachers work according to the school's educational goals	☐1	☐2	☐3	☐4	☐5	☐6
e)	I promote teaching practices based on recent educational research	☐1	☐2	☐3	☐4	☐5	☐6
f)	I praise teachers whose students are actively participating in learning	☐1	☐2	☐3	☐4	☐5	☐6
g)	When a teacher has problems in his/her classroom, I take the initiative to discuss matters	☐1	☐2	☐3	☐4	☐5	☐6
h)	I draw teachers' attention to the importance of pupils' development of critical and social capacities	☐1	☐2	☐3	☐4	☐5	☐6
i)	I pay attention to disruptive behaviour in classrooms	☐1	☐2	☐3	☐4	☐5	☐6
j)	I provide staff with opportunities to participate in school decision-making	☐1	☐2	☐3	☐4	☐5	☐6
k)	I engage teachers to help build a school culture of continuous improvement	☐1	☐2	☐3	☐4	☐5	☐6
l)	I ask teachers to participate in reviewing management practices	☐1	☐2	☐3	☐4	☐5	☐6
m)	When a teacher brings up a classroom problem, we solve the problem together	☐1	☐2	☐3	☐4	☐5	☐6
n)	I discuss the school's academic goals with teachers at faculty meetings	☐1	☐2	☐3	☐4	☐5	☐6
o)	I refer to the school's academic goals when making curricular decisions with teachers	☐1	☐2	☐3	☐4	☐5	☐6
p)	I discuss academic performance results with the faculty to identify curricular strengths and weaknesses	☐1	☐2	☐3	☐4	☐5	☐6
q)	I lead or attend in-service activities concerned with instruction	☐1	☐2	☐3	☐4	☐5	☐6
r)	I set aside time at faculty meetings for teachers to share ideas or information from in-service activities	☐1	☐2	☐3	☐4	☐5	☐6
s)	I conduct informal observations in classrooms on a regular basis (informal observations are unscheduled, last at least 5 minutes, and may or may not involve written feedback or a formal conference)	☐1	☐2	☐3	☐4	☐5	☐6
t)	I review work produced by students when evaluating classroom instruction	☐1	☐2	☐3	☐4	☐5	☐6
u)	I evaluate the performance of staff	☐1	☐2	☐3	☐4	☐5	☐6

Note: For a definition of <**the last academic year**> see Q19.

Q31 (SC35)

During the last three months, what percentage of teaching staff in your school has attended a programme of professional development with a focus on mathematics?

A programme of professional development here is a formal programme designed to enhance teaching skills or pedagogical practices. It may or may not lead to a recognised qualification. The programme must last for at least one day in total and have a focus on mathematics teaching and education.

		%
a)	All staff at your school	
b)	Staff who teach mathematics at your school	

Q32 (SC39)

Which of the following measures aimed at quality assurance and improvement do you have in your school?

(Please tick one box in each row.)

		Yes	No
a)	Written specification of the school's curricular profile and educational goals	☐1	☐2
b)	Written specification of student performance standards	☐1	☐2
c)	Systematic recording of data including teacher and student attendance and graduation rates, test results and professional development of teachers	☐1	☐2
d)	Internal evaluation/self-evaluation	☐1	☐2
e)	External evaluation	☐1	☐2
f)	Seeking written feed-back from students (e.g. regarding lessons, teachers or resources)	☐1	☐2
g)	Teacher mentoring	☐1	☐2
h)	Regular consultation aimed at school improvement with one or more experts over a period of at least six months	☐1	☐2
i)	Implementation of a standardised policy for mathematics (i.e. school curriculum with shared instructional materials accompanied by staff development and training)	☐1	☐2

Q33 (SC40)

Which of the following statements apply in your school?

A policy refers to written rules known to those concerned with the policy.

(Please tick one box in each row.)

		Yes	No
a)	The school has a policy on how to use computers in mathematics instruction (e.g. amount of computer use in mathematics lessons, use of specific mathematics computer programs)	☐1	☐2
b)	All <national modal grade for 15-year-olds> mathematics classes in the school use the same textbook	☐1	☐2
c)	Mathematics teachers in the school follow a standardised curriculum that specifies content at least on a monthly basis	☐1	☐2

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q34 (SC44)

In your school, how likely is it that a student in <national modal grade for 15-year-olds> would be transferred to another school for the following reasons?

(Please tick one box in each row.)

		Not likely	Likely	Very likely
a)	Low academic achievement	☐1	☐2	☐3
b)	High academic achievement	☐1	☐2	☐3
c)	Behavioural problems	☐1	☐2	☐3
d)	Special learning needs	☐1	☐2	☐3
e)	Parents' or guardians' request	☐1	☐2	☐3
f)	Other	☐1	☐2	☐3

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Financial education at school

The following five questions are about financial education/personal finance in your school. Financial education/personal finance involves the development of students' knowledge, confidence and skills relating to topics such as money and income; budgeting and long term planning; saving and spending; credit and debt; investment and insurance; the potential risks and benefits of financial products; and the financial landscape (including consumer rights and responsibilities and understanding of the wider financial, economic and social system).

Q35 (SC47) **Which of the statements below best describes the situation for students in <national modal grade for 15-year-olds> regarding the availability of financial education in your school?**

(Please tick only one box.)

Financial education is not available	☐1
Financial education has been available for less than two years	☐2
Financial education has been available for two years or more	☐3

Note: For a definition of <**national modal grade for 15-year-olds**> see Q5.

Q36 (SC45) **Is financial education compulsory in your school?**

(Please tick only one box.)

Yes	☐1
No	☐2

Q37 (SC46) **Which of the statements below describe the teaching of financial education in your school?**

For each statement, please indicate the number of hours of financial education of this type for students in <national modal grade for 15-year-olds> during <the last academic year>.

(Please tick as many boxes as appropriate in each row.)

		Not at all	1-4 hours a year	5-19 hours a year	20-49 hours a year	50 or more hours a year
a)	It is taught as a separate subject	☐1	☐2	☐3	☐4	☐5
b)	It is taught as a cross-curricular subject	☐1	☐2	☐3	☐4	☐5
c)	It is taught as part of <business or economics> courses	☐1	☐2	☐3	☐4	☐5
d)	It is taught as part of mathematics	☐1	☐2	☐3	☐4	☐5
e)	It is taught as part of other social sciences and humanities subjects and/or literature/language (e.g. history, geography, <home economics>, <citizenship>)	☐1	☐2	☐3	☐4	☐5
f)	It is available as an <extra-curricular activity>	☐1	☐2	☐3	☐4	☐5
g)	It is taught as part of <class teacher lessons>	☐1	☐2	☐3	☐4	☐5

Notes: For a definition of <**national modal grade for 15-year-olds**> see Q5.

For a definition of <**the last academic year**> see Q19.

<**Extra-curricular activity**> is an activity sponsored by and usually held at school but that is not part of the academic curriculum.

<**class teacher lessons**> refers to regularly scheduled time that students have with their class teacher, often for administrative or organisational purposes rather than specific subject content.

Q38 (SC52) (SC46) **Who provides financial education in your school?**

(Please tick one box in each row.)

		Yes	No
a)	Teachers	☐1	☐2
b)	People from private sector institutions (e.g. commercial bank, insurance company)	☐1	☐2
c)	People from public sector institutions (e.g. <ministry of finance>, <reserve bank>)	☐1	☐2
d)	People from non-government organisations, (e.g. <national examples>)	☐1	☐2

Q39 (SC51) **During the last twelve months, what percentage of teaching staff in your school has attended a programme of professional development with a focus on financial education?**

A programme of professional development here is a formal programme designed to enhance teaching skills or pedagogical practices. It may or may not lead to a recognised qualification. The programme must last for at least one day in total and have a focus on the teaching of financial education.

		%
a)	Staff who teach financial education in your school	
b)	All other teaching staff in your school	

Additional Question for ONLINE school questionnaire

Q40 (SC50) **Is there any final comment that you wish to make regarding any aspect of the survey (e.g. content, mode of delivery)?**

STUDENT QUESTIONNAIRE

The student questionnaire is administered after the literacy assessment and takes students about 30 minutes to complete. The core questions on home background are similar to those used in the previous PISA assessments. The questionnaire covers the following aspects:

- Student characteristics and educational career
- Family context and home resources
- Learning mathematics
- Experience with different kinds of mathematics problems at school
- Mathematics experiences
- Classroom and school climate
- Problem solving experiences

Technical terms are given in <brackets> and are adapted to the national context by the national data collection centre of the participating country or economy. In this annex, an explanation of the technical terms is given below the questionnaire item.

Student characteristics and educational career

Q1 (ST01) **What <grade> are you in?**

Grade	

Note: <**Grade**> refers to the administrative level of the student in the school. In many countries the number of years in schooling is the usual measure of grade.

Q2 (ST02) **Which one of the following <programmes> are you in?**
(Please tick only one box.)

<Programme 1>	☐ 1
<Programme 2>	☐ 2
<Programme 3>	☐ 3
<Programme 4>	☐ 4
<Programme 5>	☐ 5
<Programme 6>	☐ 6

Q3 (ST03) **On what date were you born?**
(Please write the day, month and year you were born.)

Day	*Month*	*Year*
________	________	*19* ________

Q4 (ST04) **Are you female or male?**

Female	☐ 1
Male	☐ 2

Q5 (ST05) Did you attend <ISCED 0>?

No	☐1
Yes, for one year or less	☐2
Yes, for more than one year	☐3

Note: <**ISCED 0**> refers to pre-primary education defined as the initial stage of organised instruction designed primarily to introduce very young children to a school-type environment. Instruction is school-based or centre-based, typically for children between 3 and 6 years.

Q6 (ST06) How old were you when you started <ISCED 1>?

Years	

Note: <**ISCED 1**> refers to primary education which begins between age 5 and 7 and generally lasts 4 years (e.g. Germany) to 6 years (mode of OECD countries). Primary education is designed to provide a sound basic education in reading, writing and mathematics along with an elementary understanding of other subjects.

Q7 (ST07) Have you ever repeated a <grade>?

(Please tick only one box in each row.)

		No, never	Yes, once	Yes, twice or more
a)	At <ISCED 1>	☐1	☐2	☐3
b)	At <ISCED 2>	☐1	☐2	☐3
c)	At <ISCED 3>	☐1	☐2	☐3

Notes: For a definition of <**Grade**> see Q1.

For a definition of <**ISCED 1**> level, see Q6.

<**ISCED 2**> refers to lower secondary level of education. Programmes at the start of level 2 should correspond to the point where programmes are beginning to be organised in a more subject-oriented pattern, using more specialised teachers conducting classes in their field of specialisation.

<**ISCED 3**> refers to upper secondary level of education. The final stage of secondary education in most OECD countries. Instruction is often more organised along subject-matter lines than at ISCED 2 and teachers typically need to have a higher level, or more subject-specific, qualification than at ISCED 2.

<**ISCED 3A**> is designed to provide direct access to ISCED 5A. <**ISCED 3B**> is designed to provide direct access to ISCED 5B. <**ISCED 3C**> is not designed to lead directly to ISCED 5A or 5B. These programmes lead directly to labour market, ISCED 4 programmes or other ISCED 3 programmes.

Q8 (ST08) In the last two full weeks of school, how many times did you arrive late for school?

(Please tick only one box.)

None	☐1
One or two times	☐2
Three or four times	☐3
Five or more times	☐4

Q9 (ST09) In the last two full weeks of school, how many times did you <skip> a whole school day?

(Please tick only one box.)

None	☐1
One or two times	☐2
Three or four times	☐3
Five or more times	☐4

Note: <**Skip**> refers to unauthorised failure to attend classes.

Q10 (ST115) **In the last two full weeks of school, how many times did you <skip> some classes?**

(Please tick only one box.)

None	☐1
One or two times	☐2
Three or four times	☐3
Five or more times	☐4

Note: For a definition of <**Skip**> see Q9.

Family context and home resources

Q11 (ST11) **Who usually lives at <home> with you?**

(Please tick one box in each row.)

		Yes	No
a)	Mother (including stepmother or foster mother)	☐1	☐2
b)	Father (including stepfather or foster father)	☐1	☐2
c)	Brother(s) (including stepbrothers)	☐1	☐2
d)	Sister(s) (including stepsisters)	☐1	☐2
e)	Grandparent(s)	☐1	☐2
f)	Others (e.g. cousin)	☐1	☐2

Q12a (ST12) **What is your mother's main job?**

(e.g. school teacher, kitchen-hand, sales manager)

(If she is not working now, please tell us her last main job.)

Please write in the job title:

Q12b (ST12) **What does your mother do in her main job?**

(e.g. teaches high school students, helps the cook prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work she does or did in that job:

Q13 (ST13) **What is the <highest level of schooling> completed by your mother?**

If you are not sure which box to choose, please ask the <test administrator> for help.

(Please tick only one box.)

<ISCED 3A>	☐1
<ISCED 3B, 3C>	☐2
<ISCED 2>	☐3
<ISCED 1>	☐4
She did not complete <ISCED 1>	☐5

Notes: For a definition of <**ISCED 1**> see Q6.
For a definition of <**ISCED 2**> and <**ISCED 3**> see Q7.
<**Highest level of schooling**> - should be adapted to refer to the sections of schooling that correspond to <ISCED 1> to <ISCED 3> levels.

Q14 (ST14)

Does your mother have any of the following qualifications?

If you are not sure how to answer this question, please ask the <test administrator> for help.

(Please tick one box in each row.)

		Yes	No
a)	<ISCED 6>	☐ 1	☐ 2
b)	<ISCED 5A>	☐ 1	☐ 2
c)	<ISCED 5B>	☐ 1	☐ 2
d)	<ISCED 4>	☐ 1	☐ 2

Notes: <**ISCED 6**> refers to advanced research qualification, devoted to advanced study and original research, requiring submission of a thesis or dissertation of publishable quality.

<**ISCED 5A**> refers to qualification obtained from a tertiary study programme with a strong theoretical foundation *typically* with a minimum duration of three years' full time equivalent, providing entry into a profession with high skills requirements or an advanced research programme.

<**ISCED 5B**> refers to qualification obtained in tertiary programmes that are generally more practical/technical/occupationally specific and typically shorter than ISCED 5A programmes. Typically, these programmes have a minimum of two years' full-time equivalent duration and prepare students to enter a particular occupation.

<**ISCED 4**> refers to qualification obtained in programmes that overlap the boundary between upper-secondary and post-secondary education. They are typically not significantly more advanced than programmes at ISCED 3 (see Q7) and have a full-time equivalent duration of between 6 months and 2 years.

Q15 (ST15)

What is your mother currently doing?

(Please tick only one box.)

Working full-time <for pay>	☐ 1
Working part-time <for pay>	☐ 2
Not working, but looking for a job	☐ 3
Other (e.g. home duties, retired)	☐ 4

Q16a (ST16)

What is your father's main job?

(e.g. school teacher, kitchen-hand, sales manager)

(If she is not working now, please tell us her last main job.)

Please write in the job title:

Q16b (ST16)

What does your father do in his main job?

(e.g. teaches high school students, helps the cook prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work he does or did in that job:

Q17 (ST17)

What is the <highest level of schooling> completed by your father?

If you are not sure how to answer this question, please ask the <test administrator> for help.

(Please tick only one box.)

<ISCED 3A>	☐ 1
<ISCED 3B, 3C>	☐ 2
<ISCED 2>	☐ 3
<ISCED 1>	☐ 4
He did not complete <ISCED 1>	☐ 5

Notes: For a definition of <**ISCED 1**> see Q6.

For a definition of <**ISCED 2**> and <**ISCED 3**> see Q7.

For a definition of <**Highest level of schooling**> see Q13.

Q18 (ST18)

Does your father have any of the following qualifications?

If you are not sure which box to choose, please ask the <test administrator> for help.

(Please tick one box in each row.)

		Yes	No
a)	<ISCED 6>	☐1	☐2
b)	<ISCED 5A>	☐1	☐2
c)	<ISCED 5B>	☐1	☐2
d)	<ISCED 4>	☐1	☐2

Note: For a definition of <**ISCED 4**> to <**ISCED 6**> levels see Q14.

Q19 (ST19)

What is your father currently doing?

(Please tick only one box.)

Working full-time <for pay>	☐1
Working part-time <for pay>	☐2
Not working, but looking for a job	☐3
Other (e.g. home duties, retired)	☐4

Q20 (ST20)

In what country were you and your parents born?

(Please tick one box in each column.)

	You	Mother	Father
<Country A>	☐1	☐1	☐1
<Country B>	☐2	☐2	☐2
<Country C>	☐3	☐3	☐3
<Country D>	☐4	☐4	☐4
<…etc.>	☐5	☐5	☐5
Other country	☐6	☐6	☐6

Note: Usually <**Country A**> is the country of the test. The final variable is usually '**Other country**'. Between these two variables, each country or economy may choose certain countries of origin for this question.

Q21 (ST21)

If you were NOT born in <country of test>, how old were you when you arrived in <country of test>?

If you were less than 12 months old, please write zero (0).

If you were born in <country of test> please skip this question and go to Q22.

Years	

Q22 (ST25)

What language do you speak at home most of the time?

(Please tick only one box.)

<Language 1>	☐1
<Language 2>	☐2
<Language 3>	☐3
< …etc. >	☐4
Other language	☐5

Q23 (ST26) **Which of the following are in your home?**
(Please tick one box in each row.)

		Yes	No
a)	A desk to study at	☐1	☐2
b)	A room of your own	☐1	☐2
c)	A quiet place to study	☐1	☐2
d)	A computer you can use for school work	☐1	☐2
e)	Educational software	☐1	☐2
f)	A link to the Internet	☐1	☐2
g)	Classic literature (e.g. <Shakespeare>)	☐1	☐2
h)	Books of poetry	☐1	☐2
i)	Works of art (e.g. paintings)	☐1	☐2
j)	Books to help with your school work	☐1	☐2
k)	<Technical reference books>	☐1	☐2
l)	A dictionary	☐1	☐2
m)	A dishwasher	☐1	☐2
n)	A <DVD> player	☐1	☐2
o)	<Country-specific wealth item 1>	☐1	☐2
p)	<Country-specific wealth item 2>	☐1	☐2
q)	<Country-specific wealth item 3>	☐1	☐2

Q24 (ST27) **How many of these are there at your home?**
(Please tick only one box in each row.)

		None	One	Two	Three or more
a)	Cellular phones	☐1	☐2	☐3	☐4
b)	Televisions	☐1	☐2	☐3	☐4
c)	Computers	☐1	☐2	☐3	☐4
d)	Cars	☐1	☐2	☐3	☐4
e)	Rooms with a bath or shower	☐1	☐2	☐3	☐4

Q25 (ST28) **How many books are there in your home?**
There are usually about 40 books per metre of shelving. Do not include magazines, newspapers, or your schoolbooks.
(Please tick only one box.)

0-10 books	☐1
11-25 books	☐2
26-100 books	☐3
101-200 books	☐4
201-500 books	☐5
More than 500 books	☐6

Learning mathematics

Q26 (ST29) **Thinking about your views on mathematics: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	I enjoy reading about mathematics	☐1	☐2	☐3	☐4
b)	Making an effort in mathematics is worth it because it will help me in the work that I want to do later on	☐1	☐2	☐3	☐4
c)	I look forward to my mathematics lessons	☐1	☐2	☐3	☐4
d)	I do mathematics because I enjoy it	☐1	☐2	☐3	☐4
e)	Learning mathematics is worthwhile for me because it will improve my career <prospects, chances>	☐1	☐2	☐3	☐4
f)	I am interested in the things I learn in mathematics	☐1	☐2	☐3	☐4
g)	Mathematics is an important subject for me because I need it for what I want to study later on	☐1	☐2	☐3	☐4
h)	I will learn many things in mathematics that will help me get a job	☐1	☐2	☐3	☐4

Q27 (ST35) **Thinking about how people important to you view mathematics: how strongly do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Most of my friends do well in mathematics	☐1	☐2	☐3	☐4
b)	Most of my friends work hard at mathematics	☐1	☐2	☐3	☐4
c)	My friends enjoy taking mathematics tests	☐1	☐2	☐3	☐4
d)	My parents believe it's important for me to study mathematics	☐1	☐2	☐3	☐4
e)	My parents believe that mathematics is important for my career	☐1	☐2	☐3	☐4
f)	My parents like mathematics	☐1	☐2	☐3	☐4

Q28 (ST37) **How confident do you feel about having to do the following mathematics tasks?**
(Please tick only one box in each row.)

		Very confident	Confident	Not very confident	Not at all confident
a)	Using a <train timetable> to work out how long it would take to get from one place to another	☐1	☐2	☐3	☐4
b)	Calculating how much cheaper a TV would be after a 30% discount	☐1	☐2	☐3	☐4
c)	Calculating how many square metres of tiles you need to cover a floor	☐1	☐2	☐3	☐4
d)	Understanding graphs presented in newspapers	☐1	☐2	☐3	☐4
e)	Solving an equation like 3x+5= 17	☐1	☐2	☐3	☐4
f)	Finding the actual distance between two places on a map with a 1:10,000 scale	☐1	☐2	☐3	☐4
g)	Solving an equation like 2(x+3) = (x + 3) (x - 3)	☐1	☐2	☐3	☐4
h)	Calculating the petrol consumption rate of a car	☐1	☐2	☐3	☐4

Q29 (ST42) **Thinking about studying mathematics: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	I often worry that it will be difficult for me in mathematics classes	☐1	☐2	☐3	☐4
b)	I am just not good at mathematics	☐1	☐2	☐3	☐4
c)	I get very tense when I have to do mathematics homework	☐1	☐2	☐3	☐4
d)	I get good <grades> in mathematics	☐1	☐2	☐3	☐4
e)	I get very nervous doing mathematics problems	☐1	☐2	☐3	☐4
f)	I learn mathematics quickly	☐1	☐2	☐3	☐4
g)	I have always believed that mathematics is one of my best subjects	☐1	☐2	☐3	☐4
h)	I feel helpless when doing a mathematics problem	☐1	☐2	☐3	☐4
i)	In my mathematics class, I understand even the most difficult work	☐1	☐2	☐3	☐4
j)	I worry that I will get poor <grades> in mathematics	☐1	☐2	☐3	☐4

Note: <**Grades**> refers to the teacher's standardised evaluation of student performance in a course or subject and should be adapted to the national context.

Q30 (ST43) **Thinking about your mathematics lessons: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	If I put in enough effort I can succeed in mathematics	☐1	☐2	☐3	☐4
b)	Whether or not I do well in mathematics is completely up to me	☐1	☐2	☐3	☐4
c)	Family demands or other problems prevent me from putting a lot of time into my mathematics work	☐1	☐2	☐3	☐4
d)	If I had different teachers, I would try harder in mathematics	☐1	☐2	☐3	☐4
e)	If I wanted to, I could do well in mathematics	☐1	☐2	☐3	☐4
f)	I do badly in mathematics whether or not I study for my exams	☐1	☐2	☐3	☐4

Q31 (ST44) **Suppose that you are a student in the following situation:**
Each week, your mathematics teacher gives a short quiz. Recently you have done badly on these quizzes. Today you are trying to figure out why.
How likely are you to have these thoughts or feelings in this situation?
(Please tick only one box in each row.)

		Very likely	Likely	Slightly likely	Not at all likely
a)	I'm not very good at solving mathematics problems	☐1	☐2	☐3	☐4
b)	My teacher did not explain the concepts well this week	☐1	☐2	☐3	☐4
c)	This week I made bad guesses on the quiz	☐1	☐2	☐3	☐4
d)	Sometimes the course material is too hard	☐1	☐2	☐3	☐4
e)	The teacher did not get students interested in the material	☐1	☐2	☐3	☐4
f)	Sometimes I am just unlucky	☐1	☐2	☐3	☐4

Q32 (ST46) **Thinking about the mathematics you do for school: to what extent do you agree with the following statements?**

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	I finish my homework in time for mathematics class	☐1	☐2	☐3	☐4
b)	I work hard on my mathematics homework	☐1	☐2	☐3	☐4
c)	I am prepared for my mathematics exams	☐1	☐2	☐3	☐4
d)	I study hard for mathematics quizzes	☐1	☐2	☐3	☐4
e)	I keep studying until I understand mathematics material	☐1	☐2	☐3	☐4
f)	I pay attention in mathematics class	☐1	☐2	☐3	☐4
g)	I listen in mathematics class	☐1	☐2	☐3	☐4
h)	I avoid distractions when I am studying mathematics	☐1	☐2	☐3	☐4
i)	I keep my mathematics work well organised	☐1	☐2	☐3	☐4

Q33 (ST48) **For each pair of statements, please choose the item that best describes you.**

(Please tick only one of the following two boxes.)

a)	I intend to take additional mathematics courses after school finishes	☐1
	I intend to take additional <test language> courses after school finishes	☐2
b)	I plan on majoring in a subject in <college> that requires mathematics skills	☐1
	I plan on majoring in a subject in <college> that requires <science> skills	☐2
c)	I am willing to study harder in my mathematics classes than is required	☐1
	I am willing to study harder in my <test language> classes than is required	☐2
d)	I plan on <taking> as many mathematics classes as I can during my education	☐1
	I plan on <taking> as many <science> classes as I can during my education	☐2
e)	I am planning on pursuing a career that involves a lot of mathematics	☐1
	I am planning on pursuing a career that involves a lot of <science>	☐2

Notes: <**Test language**> refers to the language of instruction in which the PISA reading assessment is administered. In some countries <test language> may be taught in different school subjects, e.g. English language and English literature. If this is the case, <test language> must be adapted accordingly.

<**Science**> refers only to the core science subjects of physics, chemistry, Earth science and biology either taught in the country's curriculum as separate science subjects, or taught within a single 'integrated-science' subject. The term does not include related subjects such as engineering, technology, mathematics, psychology, economics, nor possible Earth science topics included in geography courses. In many countries this term has been adapted to the national context.

<**College**> refers to university level or tertiary education and should be adapted to the national context.

<**Take**> refers to students who have the option of selecting additional elective courses in school. For countries with a mandatory curriculum that does not allow any choices or options, this item should be adapted to the national context.

Q34 (ST49) **How often do you do the following things at school and outside of school?**

(Please tick only one box in each row.)

		Always or almost always	Often	Sometimes	Never or rarely
a)	I talk about mathematics problems with my friends	☐1	☐2	☐3	☐4
b)	I help my friends with mathematics	☐1	☐2	☐3	☐4
c)	I do mathematics as an <extracurricular> activity	☐1	☐2	☐3	☐4
d)	I take part in mathematics competitions	☐1	☐2	☐3	☐4
e)	I do mathematics more than 2 hours a day outside of school	☐1	☐2	☐3	☐4
f)	I play chess	☐1	☐2	☐3	☐4
g)	I program computers	☐1	☐2	☐3	☐4
h)	I participate in a mathematics club	☐1	☐2	☐3	☐4

Note: <**Extracurricular**> refers to activities performed by students that fall outside the realm of the normal curriculum of school education.

Q35 (ST53) **For each group of three items, please choose the item that best describes your approach to mathematics**

(Please tick only one of the following three boxes.)

a)	When I study for a mathematics test, I try to work out what the most important parts to learn are	☐1
	When I study for a mathematics test, I try to understand new concepts by relating them to things I already know	☐2
	When I study for a mathematics test, I learn as much as I can off by heart	☐3
b)	When I study mathematics, I try to figure out which concepts I still have not understood properly	☐1
	When I study mathematics, I think of new ways to get the answer	☐2
	When I study mathematics, I make myself check to see if I remember the work I have already done	☐3
c)	When I study mathematics, I try to relate the work to things I have learnt in other subjects	☐1
	When I study mathematics, I start by working out exactly what I need to learn	☐2
	When I study mathematics, I go over some problems so often that I feel as if I could solve them in my sleep	☐3
d)	In order to remember the method for solving a mathematics problem, I go through examples again and again	☐1
	I think about how the mathematics I have learnt can be used in everyday life	☐2
	When I cannot understand something in mathematics, I always search for more information to clarify the problem	☐3

Q36 (ST55) **How many hours do you typically spend per week attending <out-of-school-time lessons> in the following subjects?**

These are only lessons in subjects that you are also learning at school, which you spend learning extra time on outside of normal school hours. The lessons may be given at your school, at your home or somewhere else.

(Please tick only one box in each row.)

		I do not attend <out-of-school-time lessons> in this subject	Less than 2 hours a week	2 or more but less than 4 hours a week	4 or more but less than 6 hours a week	6 or more hours a week
a)	<Test language>	☐1	☐2	☐3	☐4	☐5
b)	Mathematics	☐1	☐2	☐3	☐4	☐5
c)	<Science>	☐1	☐2	☐3	☐4	☐5
d)	Other Subjects	☐1	☐2	☐3	☐4	☐5

Notes: <**Out-of-school-time lessons**> refers to any lessons in the student's school subjects that he or she spends extra time learning outside of normal school hours. The lessons might be held at school, at home, or elsewhere.

For a definition of <**test language**> and <**science**> see Q33.

Q37 (ST57) **Thinking about all school subjects: on average, how many hours do you spend each week on the following?**
When answering, include time spent on the weekend too.

		Hours per week
a)	Homework or other study set by your teachers	
b)	Out of the time spent in (a), how many hours do you work on your homework with somebody overlooking and providing help if necessary (“guided homework”), either at school or elsewhere?	
c)	Work with a personal <tutor> (whether paid or not)	
d)	Attend <out of school> classes organised by a commercial company, and paid for by your parents	
e)	Study with a parent or other family member	
f)	Repeat and train content from school lessons by working on a computer (e.g. learn vocabulary with training software)	

Note: <**Out of school**> refers to any classes in the student’s school subjects that he or she spends extra time learning outside of normal school hours. The lessons might be held at school, at home, or elsewhere.

Q38 (ST61) **How often have you encountered the following types of mathematics tasks during your time at school?**
(Please tick only one box on each row.)

		Frequently	Sometimes	Rarely	Never
a)	Working out from a <train timetable> how long it would take to get from one place to another	☐ 1	☐ 2	☐ 3	☐ 4
b)	Calculating how much more expensive a computer would be after adding tax	☐ 1	☐ 2	☐ 3	☐ 4
c)	Calculating how many square metres of tiles you need to cover a floor	☐ 1	☐ 2	☐ 3	☐ 4
d)	Understanding scientific tables presented in an article	☐ 1	☐ 2	☐ 3	☐ 4
e)	Solving an equation like $6x^2 + 5 = 29$	☐ 1	☐ 2	☐ 3	☐ 4
f)	Finding the actual distance between two places on a map with a 1:10,000 scale	☐ 1	☐ 2	☐ 3	☐ 4
g)	Solving an equation like $2(x+3) = (x + 3)(x - 3)$	☐ 1	☐ 2	☐ 3	☐ 4
h)	Calculating the power consumption of an electronic appliance per week	☐ 1	☐ 2	☐ 3	☐ 4
i)	Solving an equation like $3x+5=17$	☐ 1	☐ 2	☐ 3	☐ 4

Q39 (ST62) **Thinking about mathematical concepts: how familiar are you with the following terms?**
(Please tick only one box in each row.)

		Never heard of it	Heard of it once or twice	Heard of it a few times	Heard of it often	Know it well, understand the concept
a)	Exponential Function	☐1	☐2	☐3	☐4	☐5
b)	Divisor	☐1	☐2	☐3	☐4	☐5
c)	Quadratic Function	☐1	☐2	☐3	☐4	☐5
d)	<Proper Number>	☐1	☐2	☐3	☐4	☐5
e)	Linear Equation	☐1	☐2	☐3	☐4	☐5
f)	Vectors	☐1	☐2	☐3	☐4	☐5
g)	Complex Number	☐1	☐2	☐3	☐4	☐5
h)	Rational Number	☐1	☐2	☐3	☐4	☐5
i)	Radicals	☐1	☐2	☐3	☐4	☐5
j)	<Subjunctive Scaling>	☐1	☐2	☐3	☐4	☐5
k)	Polygon	☐1	☐2	☐3	☐4	☐5
l)	<Declarative Fraction>	☐1	☐2	☐3	☐4	☐5
m)	Congruent Figure	☐1	☐2	☐3	☐4	☐5
n)	Cosine	☐1	☐2	☐3	☐4	☐5
o)	Arithmetic Mean	☐1	☐2	☐3	☐4	☐5
p)	Probability	☐1	☐2	☐3	☐4	☐5

Q40 (ST69) **How many minutes, on average, are there in a <class period> for the following subjects?**

		Minutes
a)	Minutes in a <class period> in < test language>	
b)	Minutes in a <class period> in mathematics	
c)	Minutes in a <class period> in <science>	

Notes: <**Class period**> refers to the length of time each lesson runs for in a normal school week.
For a definition of <**test language**> and <**science**> see Q33.

Q41 (ST70) **How many <class periods> per week do you typically have for the following subjects?**

		<class periods>
a)	Number of <class periods> per week in <test language>	
b)	Number of <class periods> per week in mathematics	
c)	Number of <class periods> per week in <science>	

Notes: For a definition of <**class period**> see Q40.
For a definition of <**test language**> and <**science**> see Q33.

Q42 (ST71)	In a normal, full week at school, how many <class periods> do you have <in total>?	
		<class periods>
	Number of ALL <class periods>	

Note: For a definition of <**class period**> see Q40.

Q43 (ST72)	On average, about how many students attend your <test language> class?	
	Students	

Note: For a definition of <**test language**> see Q33.

Experience with different kinds of mathematics problems at school

Q44 (ST73)

In the box is a series of problems. Each requires you to understand a problem written in text and perform the appropriate calculations. Usually the problem talks about practical situations, but the numbers and people and places mentioned are made up. All the information you need is given. Here are two examples:

1) <Ann> is two years older than <Betty> and <Betty> is four times as old as <Sam>. When <Betty> is 30, how old is <Sam>?

2) Mr <Smith> bought a television and a bed. The television cost <$625> but he got a 10% discount. The bed cost <$200>. He paid <$20> for delivery. How much money did Mr <Smith> spend?

We want to know about your experience with these types of word problems at school. Do not solve them!

(Please tick only one box in each row.)

		Frequently	Sometimes	Rarely	Never
a)	How often have you encountered these types of problems in your **mathematics lessons**?	☐1	☐2	☐3	☐4
b)	How often have you encountered these types of problems in the **tests you have taken at school**?	☐1	☐2	☐3	☐4

Q45 (ST74)

Below are examples of another set of mathematical skills.

1) Solve 2x + 3 = 7.

2) Find the volume of a box with sides 3m, 4m and 5m.

We want to know about your experience with these types of word problems at school.

Do not solve them!

(Please tick only one box in each row.)

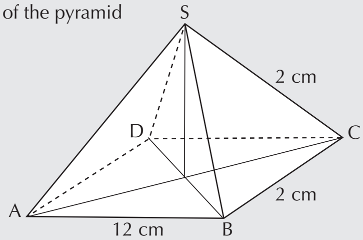
		Frequently	Sometimes	Rarely	Never
a)	How often have you encountered these types of problems in your **mathematics lessons**?	☐1	☐2	☐3	☐4
b)	How often have you encountered these types of problems in the **tests you have taken at school**?	☐1	☐2	☐3	☐4

Q46
(ST75)

In the next type of problem, you have to use mathematical knowledge and draw conclusions. There is no practical application provided. Here are two examples.

1) Here you need to use geometrical theorems:

Determine the height of the pyramid



2) Here you have to know what a prime number is:

If n is any number: can $(n+1)^2$ be a prime number?

We want to know about your experience with these types of problems at school.

Do not solve them!

(Please tick only one box in each row.)

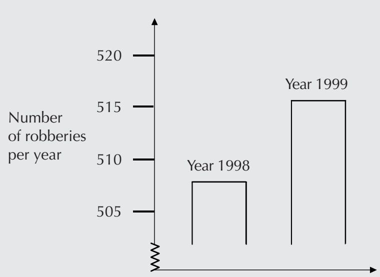
		Frequently	Sometimes	Rarely	Never
a)	How often have you encountered these types of problems in your **mathematics lessons**?	☐1	☐2	☐3	☐4
b)	How often have you encountered these types of problems in the **tests you have taken at school**?	☐1	☐2	☐3	☐4

Q47
(ST76)

In this type of problem, you have to apply suitable mathematical knowledge to find a useful answer to a problem that arises in everyday life or work. The data and information are about real situations. Here are two examples.

Example 1:

A TV reporter says "This graph shows that there is a huge increase in the number of robberies from 1998 to 1999."



Do you consider the reporter's statement to be a reasonable interpretation of the graph?
Give an explanation to support your answer.

Example 2:

For years the relationship between a person's recommended maximum heart rate and the person's age was described by the following formula:

Recommended maximum heart rate = 220 - age

Recent research showed that this formula should be modified slightly. The new formula is as follows:

Recommended maximum heart rate = 208 – (0.7 × age)

From which age onwards does the recommended maximum heart rate increase as a result of the introduction of the new formula? Show your work.

We want to know about your experience with these types of problems at school.

Do not solve them!

(Please tick only one box in each row.)

		Frequently	Sometimes	Rarely	Never
a)	How often have you encountered these types of problems in your **mathematics lessons**?	☐1	☐2	☐3	☐4
b)	How often have you encountered these types of problems in the **tests you have taken at school**?	☐1	☐2	☐3	☐4

Mathematics experiences

Q48 (ST77) **How often do these things happen in your mathematics lessons?**
(Please tick only one box in each row.)

		Every lesson	Most lessons	Some lessons	Never or hardly ever
a)	The teacher shows an interest in every student's learning	☐ 1	☐ 2	☐ 3	☐ 4
b)	The teacher gives extra help when students need it	☐ 1	☐ 2	☐ 3	☐ 4
c)	The teacher helps students with their learning	☐ 1	☐ 2	☐ 3	☐ 4
d)	The teacher continues teaching until the students understand	☐ 1	☐ 2	☐ 3	☐ 4
e)	The teacher gives students an opportunity to express opinions	☐ 1	☐ 2	☐ 3	☐ 4

Q49 (ST79) **How often do these things happen in your mathematics lessons?**
(Please tick only one box in each row.)

		Every lesson	Most lessons	Some lessons	Never or hardly ever
a)	The teacher sets clear goals for our learning	☐ 1	☐ 2	☐ 3	☐ 4
b)	The teacher asks me or my classmates to present our thinking or reasoning at some length	☐ 1	☐ 2	☐ 3	☐ 4
c)	The teacher gives different work to classmates who have difficulties learning and/or to those who can advance faster	☐ 1	☐ 2	☐ 3	☐ 4
d)	The teacher assigns projects that require at least one week to complete	☐ 1	☐ 2	☐ 3	☐ 4
e)	The teacher tells me about how well I am doing in my mathematics class	☐ 1	☐ 2	☐ 3	☐ 4
f)	The teacher asks questions to check whether we have understood what was taught	☐ 1	☐ 2	☐ 3	☐ 4
g)	The teacher has us work in small groups to come up with joint solutions to a problem or task	☐ 1	☐ 2	☐ 3	☐ 4
h)	At the beginning of a lesson, the teacher presents a short summary of the previous lesson	☐ 1	☐ 2	☐ 3	☐ 4
i)	The teacher asks us to help plan classroom activities or topics	☐ 1	☐ 2	☐ 3	☐ 4
j)	The teacher gives me feedback on my strengths and weaknesses in mathematics	☐ 1	☐ 2	☐ 3	☐ 4
k)	The teacher tells us what is expected of us when we get a test, quiz or assignment	☐ 1	☐ 2	☐ 3	☐ 4
l)	The teacher tells us what we have to learn	☐ 1	☐ 2	☐ 3	☐ 4
m)	The teacher tells me what I need to do to become better in mathematics	☐ 1	☐ 2	☐ 3	☐ 4

Q50 (ST80)

Thinking about the mathematics teacher that taught your last mathematics class: How often does each of the following happen?

(Please tick only one box in each row.)

		Always or almost always	Often	Sometimes	Never or rarely
a)	The teacher asks questions that make us reflect on the problem	☐ 1	☐ 2	☐ 3	☐ 4
b)	The teacher gives problems that require us to think for an extended time	☐ 1	☐ 2	☐ 3	☐ 4
c)	The teacher asks us to decide on our own procedures for solving complex problems	☐ 1	☐ 2	☐ 3	☐ 4
d)	The teacher presents problems for which there is no immediately obvious method of solution	☐ 1	☐ 2	☐ 3	☐ 4
e)	The teacher presents problems in different contexts so that students know whether they have understood the concepts	☐ 1	☐ 2	☐ 3	☐ 4
f)	The teacher helps us to learn from mistakes we have made	☐ 1	☐ 2	☐ 3	☐ 4
g)	The teacher asks us to explain how we have solved a problem	☐ 1	☐ 2	☐ 3	☐ 4
h)	The teacher presents problems that require students to apply what they have learnt to new contexts	☐ 1	☐ 2	☐ 3	☐ 4
i)	The teacher gives problems that can be solved in several different ways	☐ 1	☐ 2	☐ 3	☐ 4

Q51 (ST81)

How often do these things happen in your mathematics lessons?

(Please tick only one box in each row.)

		Every lesson	Most lessons	Some lessons	Never or hardly ever
a)	Students don't listen to what the teacher says	☐ 1	☐ 2	☐ 3	☐ 4
b)	There is noise and disorder	☐ 1	☐ 2	☐ 3	☐ 4
c)	The teacher has to wait a long time for students to <quiet down>	☐ 1	☐ 2	☐ 3	☐ 4
d)	Students cannot work well	☐ 1	☐ 2	☐ 3	☐ 4
e)	Students don't start working for a long time after the lesson begins	☐ 1	☐ 2	☐ 3	☐ 4

Q52 (ST82)

Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers, then let us know to what extent you agree with the final statement.

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Ms. <name> sets mathematics homework every other day. She always gets the answers back to students before examinations **Ms. <name> is concerned about her students' learning**	☐ 1	☐ 2	☐ 3	☐ 4
b)	Mr. <name> sets mathematics homework once a week. He always gets the answers back to students before examinations **Mr. <name> is concerned about his students' learning**	☐ 1	☐ 2	☐ 3	☐ 4
c)	Ms. <name> sets mathematics homework once a week. She never gets the answers back to students before examinations **Ms. <name> is concerned about her students' learning**	☐ 1	☐ 2	☐ 3	☐ 4

Q53 (ST83) **Thinking about the mathematics teacher who taught your last mathematics class: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	My teacher lets us know we need to work hard	☐ 1	☐ 2	☐ 3	☐ 4
b)	My teacher provides extra help when needed	☐ 1	☐ 2	☐ 3	☐ 4
c)	My teacher helps students with their learning	☐ 1	☐ 2	☐ 3	☐ 4
d)	My teacher gives students the opportunity to express opinions	☐ 1	☐ 2	☐ 3	☐ 4

Q54 (ST84) **Below you will find descriptions of three mathematics teachers. Read each of the descriptions of these teachers, then let us know to what extent you agree with the final statement.**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	The students' in Ms. <name's> class frequently interrupt her lessons. She always arrives five minutes early to class **Ms. <name> is in control of her classroom**	☐ 1	☐ 2	☐ 3	☐ 4
b)	The students' in Ms. <name's> class are calm and orderly She always arrives on time to class **Ms. <name> is in control of her classroom**	☐ 1	☐ 2	☐ 3	☐ 4
c)	The students' in Mr. <name's> class frequently interrupt his lessons. As a result, he often arrives five minutes late to class **Mr. <name> is in control of his classroom**	☐ 1	☐ 2	☐ 3	☐ 4

Q55 (ST85) **Thinking about the mathematics teacher who taught your last mathematics class: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	My teacher gets students to listen to him or her	☐ 1	☐ 2	☐ 3	☐ 4
b)	My teacher keeps the class orderly	☐ 1	☐ 2	☐ 3	☐ 4
c)	My teacher starts lessons on time	☐ 1	☐ 2	☐ 3	☐ 4
d)	The teacher has to wait a long time for students to <quiet down>	☐ 1	☐ 2	☐ 3	☐ 4

Classroom and school climate

Q56 (ST86) **Thinking about the teachers at your school: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Students get along well with most teachers	☐ 1	☐ 2	☐ 3	☐ 4
b)	Most teachers are interested in students' well-being	☐ 1	☐ 2	☐ 3	☐ 4
c)	Most of my teachers really listen to what I have to say	☐ 1	☐ 2	☐ 3	☐ 4
d)	If I need extra help, I will receive it from my teachers	☐ 1	☐ 2	☐ 3	☐ 4
e)	Most of my teachers treat me fairly	☐ 1	☐ 2	☐ 3	☐ 4

Q57 (ST87) Thinking about your school: to what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	I feel like an outsider (or left out of things) at school	☐1	☐2	☐3	☐4
b)	I make friends easily at school	☐1	☐2	☐3	☐4
c)	I feel like I belong at school	☐1	☐2	☐3	☐4
d)	I feel awkward and out of place in my school	☐1	☐2	☐3	☐4
e)	Other students seem to like me	☐1	☐2	☐3	☐4
f)	I feel lonely at school	☐1	☐2	☐3	☐4
g)	I feel happy at school	☐1	☐2	☐3	☐4
h)	Things are ideal in my school	☐1	☐2	☐3	☐4
i)	I am satisfied with my school	☐1	☐2	☐3	☐4

Q58 (ST88) Thinking about what you have learnt at school: to what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	School has done little to prepare me for adult life when I leave school	☐1	☐2	☐3	☐4
b)	School has been a waste of time	☐1	☐2	☐3	☐4
c)	School has helped give me confidence to make decisions	☐1	☐2	☐3	☐4
d)	School has taught me things which could be useful in a job	☐1	☐2	☐3	☐4

Q59 (ST89) Thinking about your school: to what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Trying hard at school will help me get a good job	☐1	☐2	☐3	☐4
b)	Trying hard at school will help me get into a good <college>	☐1	☐2	☐3	☐4
c)	I enjoy receiving good <grades>	☐1	☐2	☐3	☐4
d)	Trying hard at school is important	☐1	☐2	☐3	☐4

Notes: For a definition of <**college**> see Q33.
For a definition of <**grades**> see Q29.

Q60 (ST91) Thinking about your school: to what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	If I put in enough effort, I can succeed in school	☐1	☐2	☐3	☐4
b)	It is completely my choice whether or not I do well at school	☐1	☐2	☐3	☐4
c)	Family demands or other problems prevent me from putting a lot of time into my school work	☐1	☐2	☐3	☐4
d)	If I had different teachers, I would try harder at school	☐1	☐2	☐3	☐4
e)	If I wanted to, I could perform well at school	☐1	☐2	☐3	☐4
f)	I perform poorly at school whether or not I study for my exams	☐1	☐2	☐3	☐4

Problem solving experiences

Q61 (ST93) **How well does each of the following statements below describe you?**
(Please tick only one box in each row.)

		Very much like me	Mostly like me	Somewhat like me	Not much like me	Not at all like me
a)	When confronted with a problem, I give up easily	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b)	I put off difficult problems	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c)	I remain interested in the tasks that I start	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d)	I continue working on tasks until everything is perfect	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
e)	When confronted with a problem, I do more than what is expected of me	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q62 (ST94) **How well does each of the following statements below describe you?**
(Please tick only one box in each row.)

		Very much like me	Mostly like me	Somewhat like me	Not much like me	Not at all like me
a)	I can handle a lot of information	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b)	I am quick to understand things	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c)	I seek explanations for things	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d)	I can easily link facts together	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
e)	I like to solve complex problems	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q63 (ST96) *Suppose that you have been sending text messages from your mobile phone for several weeks. Today, however, you can't send text messages. You want to try to solve the problem.*

What would you do? For each suggestion, tick the option that best applies to you.
(Please tick only one box in each row.)

		I would definitely do this	I would probably do this	I would probably not do this	I would definitely not do this
a)	I press every button possible to find out what is wrong	☐ 1	☐ 2	☐ 3	☐ 4
b)	I think about what might have caused the problem and what I can do to solve it	☐ 1	☐ 2	☐ 3	☐ 4
c)	I read the manual	☐ 1	☐ 2	☐ 3	☐ 4
d)	I ask a friend for help	☐ 1	☐ 2	☐ 3	☐ 4

Q64 (ST101) *Suppose that you are planning a trip to the zoo with your brother. You don't know which route to take to get there.*

What would you do? For each suggestion, tick the option that best applies to you.
(Please tick only one box in each row.)

		I would definitely do this	I would probably do this	I would probably not do this	I would definitely not do this
a)	I read the zoo brochure to see if it says how to get there	☐ 1	☐ 2	☐ 3	☐ 4
b)	I study a map and work out the best route	☐ 1	☐ 2	☐ 3	☐ 4
c)	I leave it to my brother to worry about how to get there	☐ 1	☐ 2	☐ 3	☐ 4
d)	I know roughly where it is, so I suggest we just start driving	☐ 1	☐ 2	☐ 3	☐ 4

Q65 (ST104)

Suppose that you arrive at the train station. There is a ticket machine that you have never used before. You want to buy a ticket.

What would you do? For each suggestion, tick the option that best applies to you.

(Please tick only one box in each row.)

		I would definitely do this	I would probably do this	I would probably not do this	I would definitely not do this
a)	I check how similar it is to other ticket machines I have used	☐1	☐2	☐3	☐4
b)	I try out all the buttons to see what happens	☐1	☐2	☐3	☐4
c)	I ask someone for help	☐1	☐2	☐3	☐4
d)	I try to find a ticket office at the station to buy a ticket	☐1	☐2	☐3	☐4

EDUCATIONAL CAREER QUESTIONNAIRE

As in previous surveys, additional questionnaire materials were developed and offered as international options to the participating countries and economies. In PISA 2012, these are the educational career questionnaire and the ICT familiarity questionnaire for students, as well as the parent questionnaire.

The educational career questionnaire covers the following aspects:

- Educational career
- Preparation for future careers
- Support with language learning

Educational career

Q1 (EC01) **Did you ever miss two or more consecutive months of <ISCED 1>?**
(Please tick only one box.)

No, never	☐ 1
Yes, once	☐ 2
Yes, twice or more	☐ 3

Note: For a definition of <**ISCED 1**> see the note in Q6 of the student questionnaire.

Q2 (EC02) **Did you ever miss two or more consecutive months of <ISCED 2>?**
(Please tick only one box.)

No, never	☐ 1
Yes, once	☐ 2
Yes, twice or more	☐ 3

Note: For a definition of <**ISCED 2**> see the note in Q7 of the student questionnaire.

Preparation for future careers

Q3 (EC03) **Have you done any of the following to find out about future study or types of work?**
(Please tick one box in each row.)

		Yes	No, never
a)	I did an internship	☐ 1	☐ 2
b)	I attended <job shadowing or work-site visits>	☐ 1	☐ 2
c)	I visited a <job fair>	☐ 1	☐ 2
d)	I spoke to a <career advisor> at my school	☐ 1	☐ 2
e)	I spoke to a <career advisor> outside of my school	☐ 1	☐ 2
f)	I completed a questionnaire to find out about my interests and abilities	☐ 1	☐ 2
g)	I researched the internet for information about careers	☐ 1	☐ 2
h)	I went on an organised tour in an <ISCED 3-5> institution	☐ 1	☐ 2
i)	I researched the internet for information about <ISCED 3-5> programmes	☐ 1	☐ 2
j)	<country specific item>	☐ 1	☐ 2

Notes: A '**tour in an <ISCED 3-5> institution**' is an organised visit to a college, a university or any other institution in upper secondary or tertiary education that aims at informing prospective students about studying at upper levels in general and about the specific offers of that institution.

For a definition of <**ISCED 3**> see the note in Q7 of the student questionnaire.

For a definition of <**ISCED 4**> and <**ISCED 5**> see the note in Q14 of the student questionnaire.

Q4 (EC04) **Which of the following skills have you acquired?**
(Please tick all that apply.)

		Yes, at school	Yes, out of school	No, never
a)	How to find information on jobs I am interested in	☐ 1	☐ 2	☐ 3
b)	How to search for a job	☐ 1	☐ 2	☐ 3
c)	How to write a <résumé> or a summary of my qualifications	☐ 1	☐ 2	☐ 3
d)	How to prepare for a job interview	☐ 1	☐ 2	☐ 3
e)	How to find information on <ISCED 3-5> programs I am interested in	☐ 1	☐ 2	☐ 3
f)	How to find information on student financing (e.g. student loans or grants)	☐ 1	☐ 2	☐ 3

Notes: For a definition of <**ISCED 3**> see the note in Q7 of the student questionnaire.

For a definition of <**ISCED 4**> and <**ISCED 5**> see the note in Q14 of the student questionnaire.

Support with language learning

Q5 (EC05) **What is the first language you learned at home?**
(Please tick only one box.)

The first language I learned at home was <test language> or <other official national language(s) or dialect(s)>	☐ 1	*Please stop here. Thank you for your participation.*
I learned <test language> or <other official national language(s) or dialect(s)> and another language at the same time at home	☐ 2	*Go to Q6.*
The first language I learned at home was a language other than <test language> or <other official national language(s) or dialect(s)>	☐ 3	*Go to Q6.*

Notes: <**Test language**> refers to the name of the language used in the PISA mathematics literacy test.

<**Other official national language or dialect**>, if there is no other official language in a country, this category should be omitted. If there are more than one official languages or dialects in a country it should be replaced by the name of the respective second (and third) official language or dialect.

Q6 (EC06) **How old were you when you started learning <test language>?**

(Please tick only one box.)

0 to 3 years old	☐1
4 to 6 years old	☐2
7 to 9 years old	☐3
10 to 12 years old	☐4
13 years or older	☐5

Note: For a definition of <**test language**> see Q5.

Q7 (EC07) **Which language do you usually speak with the following people?**

(Please tick only one box in each row.)

		Mostly my <heritage language>	About equally often my <heritage language> and <test language>	Mostly <test language>	Not applicable
a)	My mother	☐1	☐2	☐3	☐4
b)	My father	☐1	☐2	☐3	☐4
c)	My brother(s) and/or sister(s)	☐1	☐2	☐3	☐4
d)	My best friend	☐1	☐2	☐3	☐4
e)	My schoolmates	☐1	☐2	☐3	☐4

Notes: For a definition of <**test language**> see Q5.

<**Heritage language**> refers to the language learned by the student at home which is not the test language. It is often also referred to as the 'mother-tongue'.

Q8 (EC08) **In which language do you usually do the following activities?**

(Please tick only one box in each row.)

		Mostly my <heritage language>	About equally often my <heritage language> and <test language>	Mostly <test language>	Not applicable
a)	Reading books, magazines or newspapers	☐1	☐2	☐3	☐4
b)	Watching TV or movies	☐1	☐2	☐3	☐4
c)	Surfing the internet	☐1	☐2	☐3	☐4
d)	Writing emails or letters	☐1	☐2	☐3	☐4

Notes: For a definition of <**test language**> see Q5.

For a definition of <**heritage language**> see Q7.

Q9 (EC09) **Have you ever attended any of the following types of lessons at school or out of school?**
(Please tick only one box in each row.)

		Yes	No, never
a)	I have attended <remedial lessons> in <test language>	☐1	☐2

Q10 (EC11) **Have you ever attended any of the following types of lessons at school or out of school?**
(Please tick only one box in each row.)

		Yes	No, never
b)	I have attended lessons in my <heritage language> (with a focus on reading, writing, or grammar)	☐1	☐2
c)	I have attended instruction in school subjects through my <heritage language> (with a focus on subject content)	☐1	☐2

Notes: For a definition of <**test language**> see Q5.
For a definition of <**heritage language**> see Q7.
<**Remedial lessons**> refers to any lesson offered outside of normal class time to help students who have fallen behind the performance level of their peers to catch up.

Q11 (EC10) **<This academic year>, how many hours a week do you receive systematic support for improving your <test language> skills at school or out of school?**
(Please tick only one box.)

None	☐1
Less than 2 hours a week	☐2
2 or more but less than 4 hours a week	☐3
4 or more but less than 6 hours a week	☐4
6 or more hours a week	☐5

Notes: For a definition of <**test language**> see Q5.
<**This academic year**> refers to the year of schooling which is not necessarily the calendar year.

Q12 (EC12) **<This academic year>, how many hours per week do you typically spend attending either lessons in your <heritage language> or instruction through your <heritage language> at school or out of school?**
(Please tick only one box.)

None	☐1
Less than 2 hours a week	☐2
2 or more but less than 4 hours a week	☐3
4 or more but less than 6 hours a week	☐4
6 or more hours a week	☐5

Notes: For a definition of <**This academic year**> see Q11.
For a definition of <**heritage language**> see Q7.

Q13 (ST22) **Was your mother born in <country of test>?**
(Please tick only one box.)

No	☐1
Yes	☐2

If YES please stop here. Thank you for your participation.

Q14 (ST23)

Below you will find statements about <host culture> and <heritage culture>. <Host culture> refers to the culture and country in which you now live. <Heritage culture> refers to the culture and country in which your mother was born.

To what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	I like to have <host culture> friends	☐1	☐2	☐3	☐4
b)	I like to have <heritage culture> friends	☐1	☐2	☐3	☐4
c)	I like to participate in <host culture> celebrations	☐1	☐2	☐3	☐4
d)	I like to participate in <heritage culture> celebrations	☐1	☐2	☐3	☐4
e)	I spend a lot of time with <host culture> friends	☐1	☐2	☐3	☐4
f)	I spend a lot of time with <heritage culture> friends	☐1	☐2	☐3	☐4
g)	I participate in <host culture> celebrations	☐1	☐2	☐3	☐4
h)	I participate in <heritage culture> celebrations	☐1	☐2	☐3	☐4

Q15 (ST24)

The statements below are about differences between <host culture> and <heritage culture>.

To what extent do you agree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	The values of people in the <host culture> and in the <heritage culture> are the same	☐1	☐2	☐3	☐4
b)	Mothers in the <host culture> and in the <heritage culture> treat their children in the same way	☐1	☐2	☐3	☐4
c)	Pupils from the <host culture> and the <heritage culture> deal with their teachers in the same way	☐1	☐2	☐3	☐4

ICT FAMILIARITY QUESTIONNAIRE

The Information Communication Technology (ICT) familiarity questionnaire consists of questions regarding the availability of ICT, and the student's use of, and attitudes towards computers. It is administered to students after the international student questionnaire and takes about five minutes to complete.

The questionnaire covers the following aspects:

- Availability of ICT
- General computer use
- Use of ICT outside of school
- Use of ICT at school
- Attitudes toward computer

Availability of ICT

Q1 (IC01) **Are any of these devices available for you to use <u>at home</u>?**
(Please tick one box in each row.)

		Yes, and I use it	Yes, but I don't use it	No
a)	Desktop computer	$\square_1$	$\square_2$	$\square_3$
b)	Portable laptop, or notebook	$\square_1$	$\square_2$	$\square_3$
c)	<Tablet computer> (e.g. <iPad®>, <BlackBerry® PlayBook™>)	$\square_1$	$\square_2$	$\square_3$
d)	Internet connection	$\square_1$	$\square_2$	$\square_3$
e)	<Video games console>, e.g. <Sony® PlayStation® >	$\square_1$	$\square_2$	$\square_3$
f)	<Cell phone> (without Internet access)	$\square_1$	$\square_2$	$\square_3$
g)	<Cell phone> (with Internet access)	$\square_1$	$\square_2$	$\square_3$
h)	Portable music player (Mp3/Mp4 player, iPod® or similar)	$\square_1$	$\square_2$	$\square_3$
i)	Printer	$\square_1$	$\square_2$	$\square_3$
j)	USB (memory) stick	$\square_1$	$\square_2$	$\square_3$
k)	<ebook reader>, e.g. <Amazon® Kindle™>	$\square_1$	$\square_2$	$\square_3$

Q2 (IC02) **Are any of these devices available for you to use <u>at school</u>?**
(Please tick one box in each row.)

		Yes, and I use it	Yes, but I don't use it	No
a)	Desktop computer	$\square_1$	$\square_2$	$\square_3$
b)	Portable laptop or notebook	$\square_1$	$\square_2$	$\square_3$
c)	<Tablet computer> (e.g. <iPad®>, <BlackBerry® PlayBook™>)	$\square_1$	$\square_2$	$\square_3$
d)	Internet connection	$\square_1$	$\square_2$	$\square_3$
e)	Printer	$\square_1$	$\square_2$	$\square_3$
f)	USB (memory) stick	$\square_1$	$\square_2$	$\square_3$
g)	<ebook reader>, e.g. <Amazon® Kindle™>	$\square_1$	$\square_2$	$\square_3$

General computer use

Q3 (IC03) **How old were you when you first used a computer?**
*If you have **never** used a computer, please stop here. Thank you for your participation.*
(Please tick only one box.)

6 years old or younger	☐ 1
7-9 years old	☐ 2
10-12 years old	☐ 3
13 years old or older	☐ 4
I have never used a computer	☐ 5

Q4 (IC04) **How old were you when you first accessed the Internet?**
(Please tick only one box.)

6 years old or younger	☐ 1
7-9 years old	☐ 2
10-12 years old	☐ 3
13 years old or older	☐ 4
I have never accessed the Internet	☐ 5

Q5 (IC05) **During a *typical* weekday, for how long do you use the Internet at school?**
(Please tick only one box.)

No time	☐ 1
1-30 minutes per day	☐ 2
31-60 minutes per day	☐ 3
Between 1 hour and 2 hours per day	☐ 4
Between 2 hours and 4 hours per day	☐ 5
Between 4 hours and 6 hours per day	☐ 6
More than 6 hours per day	☐ 7

Q6 (IC06) **During a *typical* weekday, for how long do you use the Internet outside of school?**
(Please tick only one box.)

No time	☐ 1
1-30 minutes per day	☐ 2
31-60 minutes per day	☐ 3
Between 1 hour and 2 hours per day	☐ 4
Between 2 hours and 4 hours per day	☐ 5
Between 4 hours and 6 hours per day	☐ 6
More than 6 hours per day	☐ 7

Q7 (IC07) **On a *typical* weekend day, for how long do you use the Internet <u>outside of school</u>?**
(Please tick only one box.)

	No time	☐1
	1-30 minutes per day	☐2
	31-60 minutes per day	☐3
	Between 1 hour and 2 hours per day	☐4
	Between 2 hours and 4 hours per day	☐5
	Between 4 hours and 6 hours per day	☐6
	More than 6 hours per day	☐7

Use of ICT outside of school

Q8 (IC08) **How often do you use a computer for the following activities <u>outside of school</u>?**
(Please tick one box in each row.)

		Never or hardly ever	Once or twice a month	Once or twice a week	Almost every day	Every day
a)	Playing one-player games	☐1	☐2	☐3	☐4	☐5
b)	Playing collaborative online games	☐1	☐2	☐3	☐4	☐5
c)	Using email	☐1	☐2	☐3	☐4	☐5
d)	<Chatting online> (e.g. <MSN®>)	☐1	☐2	☐3	☐4	☐5
e)	Participating in social networks (e.g. <facebook>, <MySpace>)	☐1	☐2	☐3	☐4	☐5
f)	Browsing the Internet for fun (such as watching videos, e.g<YouTube™>)	☐1	☐2	☐3	☐4	☐5
g)	Reading news on the Internet (e.g. current affairs)	☐1	☐2	☐3	☐4	☐5
h)	Obtaining practical information from the Internet (e.g. locations, dates of events)	☐1	☐2	☐3	☐4	☐5
i)	Downloading music, films, games or software from the Internet	☐1	☐2	☐3	☐4	☐5
j)	Uploading your own created contents for sharing (e.g. music, poetry, videos, computer programs)	☐1	☐2	☐3	☐4	☐5

Q9 (IC09) **How often do you use a computer for the following activities <u>outside of school</u>?**
(Please tick one box in each row.)

		Never or hardly ever	Once or twice a month	Once or twice a week	Almost every day	Every day
a)	Browsing the Internet for schoolwork (e.g. for preparing an essay or presentation)	☐1	☐2	☐3	☐4	☐5
b)	Using email for communication with other students about schoolwork	☐1	☐2	☐3	☐4	☐5
c)	Using email for communication with teachers and submission of homework or other schoolwork	☐1	☐2	☐3	☐4	☐5
d)	Downloading, upload or browse material from my school's website (e.g. time table or course materials)	☐1	☐2	☐3	☐4	☐5
e)	Checking the school's website for announcements, e.g. absence of teachers	☐1	☐2	☐3	☐4	☐5
f)	Doing homework on the computer	☐1	☐2	☐3	☐4	☐5
g)	Sharing school related materials with other students	☐1	☐2	☐3	☐4	☐5

Use of ICT at school

Q10 (IC10) **How often do you use a computer for the following activities <u>at school</u>?**
(Please tick one box in each row.)

		Never or hardly ever	Once or twice a month	Once or twice a week	Almost every day	Every day
a)	<Chatting on line> at school	☐1	☐2	☐3	☐4	☐5
b)	Using email at school	☐1	☐2	☐3	☐4	☐5
c)	Browsing the Internet for schoolwork	☐1	☐2	☐3	☐4	☐5
d)	Downloading, uploading or browsing material from the school's website (e.g. <intranet>)	☐1	☐2	☐3	☐4	☐5
e)	Posting my work on the school's website	☐1	☐2	☐3	☐4	☐5
f)	Playing simulations at school	☐1	☐2	☐3	☐4	☐5
g)	Practicing and drilling, such as for foreign language learning or mathematics	☐1	☐2	☐3	☐4	☐5
h)	Doing homework on a school computer	☐1	☐2	☐3	☐4	☐5
i)	Using school computers for group work and communication with other students	☐1	☐2	☐3	☐4	☐5

Q11 (IC11) **Within the last month, has a computer ever been used for the following purposes in your <u>mathematics lessons</u>?**
(Please tick one box in each row.)

		Yes, students did this	Yes, but only the teacher demonstrated this	No
a)	Drawing the graph of a function (such as $y = 4x+6$)	☐1	☐2	☐3
b)	Calculating with numbers (such as calculating 5*233/8)	☐1	☐2	☐3
c)	Constructing geometric figures (e.g. an equilateral triangle with given side lengths)	☐1	☐2	☐3
d)	Entering data in a spreadsheet (e.g. in <Excel™>)	☐1	☐2	☐3
e)	Rewriting algebraic expressions and solving equations (such as $a^2+2ab+b^2$)	☐1	☐2	☐3
f)	Drawing histograms (a graph that shows the distribution of frequencies of data)	☐1	☐2	☐3
g)	Finding out how the graph of a function like $y = ax^2$ changes depending on a	☐1	☐2	☐3

Attitudes toward computers

Q12 (IC22) **Thinking about your experience with computers: to what extent do you agree with the following statements?**
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	The computer is a very useful tool for my schoolwork	☐1	☐2	☐3	☐4
b)	Doing my homework using a computer makes it more fun	☐1	☐2	☐3	☐4
c)	The Internet is a great resource for obtaining information I can use for my school work	☐1	☐2	☐3	☐4
d)	Using the computer for learning is troublesome	☐1	☐2	☐3	☐4
e)	Since anyone can upload information to the Internet, it is in general not suitable to use it for schoolwork	☐1	☐2	☐3	☐4
f)	Information obtained from the Internet is generally too unreliable to be used for school assignments	☐1	☐2	☐3	☐4

PARENT QUESTIONNAIRE

One questionnaire is administered per student. The parent questionnaire takes about 20 minutes to complete. The parent questionnaire covers parental reports related to several aspects:

- Parental background
- Cost of educational service
- Attitudes to child's school
- Parents' involvement with school
- School choice
- Parental support for learning in the home
- Mathematics in child's career and job market
- Academic and professional expectations in mathematics
- Child's past academic performance
- Child's career interests
- Parents' migration background

Parental background

Q1 (PA01) **Who will complete this questionnaire?**
(Please tick all that apply.)

a)	Mother or other female guardian	$\square_1$
b)	Father or other male guardian	$\square_2$
c)	Other	$\square_3$
	(If other, please specify):	

Q2 (PA02) **How old are the child's parents?**
(Please tick one box in each row.)

		Younger than 36 years	36 – 40 years	41 – 45 years	46 – 50 years	51 years or older
a)	Father	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
b)	Mother	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$

Q3 (PA03) **Does the child's father have any of the following qualifications?**
(Please tick one box in each row.)

		Yes	No
a)	<ISCED 5A, 6>	$\square_1$	$\square_2$
b)	<ISCED 5B>	$\square_1$	$\square_2$
c)	<ISCED 4>	$\square_1$	$\square_2$
d)	<ISCED 3A>	$\square_1$	$\square_2$

Notes: For a definition of <**ISCED 3A**> see the note in Q7 of the student questionnaire.
For a definition of <**ISCED 4**> to <**ISCED 6**> levels see the note in Q14 of the student questionnaire.

Q4a (PA04) **What is the main job of the child's father? (e.g. school teacher, kitchen-hand, sales manager.)**

(If he is not working now, please tell us his last main job.)

Please write in the job title:

Q4b (PA04) **What does the child's father do in his main job? (e.g. teaches high school students, helps the cook prepare meals in a restaurant, manages a sales team.)**

Please use a sentence to describe the kind of work he does or did in that job:

Q5 (PA05) **Does the child's mother have any of the following qualifications?**

(Please tick one box in each row.)

		Yes	No
a)	<ISCED 5A, 6>	☐1	☐2
b)	<ISCED 5B>	☐1	☐2
c)	<ISCED 4>	☐1	☐2
d)	<ISCED 3A>	☐1	☐2

Notes: For a definition of <**ISCED 3A**> see the note in Q7 of the student questionnaire.
For a definition of <**ISCED 4**> to <**ISCED 6**> see the note in Q14 of the student questionnaire.

Q6a (PA06) **What is the main job of the child's mother? (e.g. school teacher, kitchen-hand, sales manager.)**

(If she is not working now, please tell us her last main job.)

Please write in the job title:

Q6b (PA06) **What does the child's mother do in her main job? (e.g. teaches high school students, helps the cook prepare meals in a restaurant, manages a sales team.)**

Please use a sentence to describe the kind of work she does or did in that job:

Q7 (PA07) **What is your annual household income?**

Please add together the total income, before tax, from all members of your household.

Please remember we ask you to answer questions only if you feel comfortable doing so, and that all responses are kept ***strictly confidential***.

(Please tick only one box.)

Less than <$A>	☐1
<$A> or more but less than <$B>	☐2
<$B> or more but less than <$C>	☐3
<$C> or more but less than <$D>	☐4
<$D> or more but less than <$E>	☐5
<$E> or more	☐6

Notes: <**$**> - This symbol denotes the national currency of the participating country or economy.
<**$C**> is a suitably rounded value for the national **median household income**. For all households in the country, the median income is the amount for which half of the households have an income above this amount, and half, below.
<**$A**> is a suitably rounded value equal to about half of the median household income. <**$B**> is a value at about three quarters of the median household income, <**$D**> is a value at about five quarters of the median household income, and <**$E**> is a value at about one-and-a-half times the median household income.

Cost of educational service

Q8 (PA08)

Please answer the following question thinking just of expenses related to <the student who brought this questionnaire home>.

In the last twelve months, about how much would you have paid to educational providers for <u>services</u>?

In determining this, please include any tuition fees you pay to your child's school, any other fees paid to individual teachers in the school or to other teachers for any tutoring your child receives, as well as any fees for cram school.

Do **not** *include the costs of <u>goods</u> like sports equipment, school uniforms, computers or textbooks if they are not included in a general fee (that is, if you have to buy these things separately).*

(Please tick only one box.)

Nothing	☐1
<More than $0 but less than $W>	☐2
<$W or more but less than $X>	☐3
<$X or more but less than $Y>	☐4
<$Y or more but less than $Z>	☐5
<$Z> or more	☐6

Notes: <**$**> - This symbol denotes the national currency.
<**$Z**> is a suitably rounded value representing an amount of money that could be spent on an expensive education, including tutoring etc. About 5% of parents should choose this option in a national sample.
<**$W**> is a suitably rounded value for a public education without any extra tutoring. Where public education is absolutely free, this would be set at a low nominal value but above zero (e.g. 50 or 100 Euro) to take account of incidental fees that might be paid.
After determining <**$Z**> and <**$W**>, equal categories are created between these figures by using increments of (<**$Z**> - <**$W**>)/3.

Attitudes to child's school

Q9 (PA09)

We are interested in what you think about your child's school.

How much do you agree or disagree with the following statements?

(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	Most of my child's school teachers seem competent and dedicated	☐1	☐2	☐3	☐4
b)	Standards of achievement are high in my child's school	☐1	☐2	☐3	☐4
c)	I am happy with the content taught and the instructional methods used in my child's school	☐1	☐2	☐3	☐4
d)	I am satisfied with the disciplinary atmosphere in my child's school	☐1	☐2	☐3	☐4
e)	My child's progress is carefully monitored by the school	☐1	☐2	☐3	☐4
f)	My child's school provides regular and useful information on my child's progress	☐1	☐2	☐3	☐4
g)	My child's school does a good job in educating students	☐1	☐2	☐3	☐4

Parents' involvement with school

Q10 (PA10) **During the last <academic year>, have you participated in any of the following school-related activities?**
(Please tick one box in each row.)

		Yes	No
a)	Discussed my child's behaviour with a teacher on my own initiative	☐1	☐2
b)	Discussed my child's behaviour on the initiative of one of his/her teachers	☐1	☐2
c)	Volunteered in physical activities, e.g. building maintenance, carpentry, gardening or yard work	☐1	☐2
d)	Volunteered in extra-curricular activities, e.g. book club, school play, sports, field trip	☐1	☐2
e)	Volunteered in the school library or media centre	☐1	☐2
f)	<Assisted a teacher in the school>	☐1	☐2
g)	Appeared as a guest speaker	☐1	☐2
h)	Participated in local school <government>, e.g. parent council or school management committee	☐1	☐2
i)	Discussed my child's progress with a teacher on my own initiative	☐1	☐2
j)	Discussed my child's progress on the initiative of one of their teachers	☐1	☐2
k)	Volunteered in the school <canteen>	☐1	☐2

Note: <**Academic year**> refers to the year of schooling which is not necessarily the calendar year.

School choice

Q11 (PA11) *We are interested in the options you had as parents when choosing the school your child is currently attending.*
Which of the following statements best describes the schooling available to students in your location?
(Please tick only one box.)

There are two or more other schools in this area that compete with the school my child is currently attending	☐1
There is one other school in this area that competes with the school my child is currently attending	☐2
There are no other schools in this area that compete with the school my child is currently attending	☐3

Q12 (PA12) **How important are the following reasons for choosing a school for your child?**
(Please tick only one box in each row.)

		Not important	Somewhat important	Important	Very important
a)	The school is at a short distance to home	☐1	☐2	☐3	☐4
b)	The school has a good reputation	☐1	☐2	☐3	☐4
c)	The school offers particular courses or school subjects	☐1	☐2	☐3	☐4
d)	The school adheres to a particular <religious philosophy>	☐1	☐2	☐3	☐4
e)	The school has a particular approach to <pedagogy/didactics, e.g. example>	☐1	☐2	☐3	☐4
f)	Other family members attended the school	☐1	☐2	☐3	☐4
g)	<Expenses are low> (e.g. tuition, books, room and board)	☐1	☐2	☐3	☐4
h)	The school has < financial aid> available, such as a school loan, scholarship, or grant	☐1	☐2	☐3	☐4
i)	The school has an active and pleasant school climate	☐1	☐2	☐3	☐4
j)	The academic achievements of students in the school are high	☐1	☐2	☐3	☐4
k)	There is a safe school environment	☐1	☐2	☐3	☐4

Parental support for learning in the home

Q13 (PA13) **How often do you or someone else in your home do the following things with your child?**
(Please tick only one box in each row.)

		Never or hardly ever	Once or twice a year	Once or twice a month	Once or twice a week	Every day or almost every day
a)	Discuss how well my child is doing at school	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b)	Eat <the main meal> with my child around a table	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c)	Spend time just talking to my child	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d)	Help my child with his/her mathematics homework	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
e)	Discuss how my child is performing in mathematics class	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
f)	Obtain mathematics materials (e.g., applications, software, study guides etc) for my child	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
g)	Discuss with my child how mathematics can be applied in everyday life	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Mathematics in child's career and job market

Q14 (PA14) *We are interested in what you think about the need for mathematics skills in the job market today.*
How much do you agree with the following statements?
(Please tick only one box in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
a)	It is important to have good mathematics knowledge and skills in order to get any good job in today's world	☐ 1	☐ 2	☐ 3	☐ 4
b)	Employers generally appreciate strong mathematics knowledge and skills among their employees	☐ 1	☐ 2	☐ 3	☐ 4
c)	Most jobs today require some mathematics knowledge and skills	☐ 1	☐ 2	☐ 3	☐ 4
d)	It is an advantage in the job market to have good mathematics knowledge and skills	☐ 1	☐ 2	☐ 3	☐ 4

Academic and professional expectations in mathematics

Q15 (PA15) *The following questions refer to <mathematics-related careers>. A <mathematics-related career> is one that requires studying a mathematics course at a university level.*
Examples of <mathematics-related careers> include Mathematics Teacher, Economists, Financial Analyst and Computer scientist.
<Mathematics-related careers> also include many science-related careers, such as Engineers, Weather Forecasters, and Medical doctors. All of these can also be considered as <mathematics-related careers>.
Please answer the questions below.
(Please tick only one box in each row.)

		Yes	No
a)	Does anybody in your family (including you) work in a <mathematics-related career>?	☐ 1	☐ 2
b)	Does your child show an interest in working in a <mathematics-related career>?	☐ 1	☐ 2
c)	Do you expect your child will go into a <mathematics-related career>?	☐ 1	☐ 2
d)	Has your child shown interest in studying mathematics after completing <secondary school>?	☐ 1	☐ 2
e)	Do you expect your child will study mathematics after completing <secondary school>?	☐ 1	☐ 2

Child's past academic performance

Q16 (PA18) **Has your child ever repeated a <grade>?**
(Please tick only one box in each row.)

		No, never	Yes, once	Yes, twice or more
a)	At < ISCED 1>	☐1	☐2	☐3
b)	At < ISCED 2>	☐1	☐2	☐3
c)	At < ISCED 3>	☐1	☐2	☐3

Notes: <**Grade**> refers to the administrative level of the student in the school. In many countries, the number of years in schooling is the usual measure of grade.
For a definition of <**ISCED 1**> see the note in Q6 of the student questionnaire.
For a definition of <**ISCED 2**> and <**ISCED 3**> see the note in Q7 of the student questionnaire.

Child's career interests

Q17 (PA19) **Which of the following do you <u>expect</u> your child to complete?**
(Please tick as many as apply.)

a)	<ISCED 2>	☐1
b)	<ISCED 3B or C>	☐2
c)	<ISCED 3A>	☐3
d)	<ISCED 4>	☐4
e)	<ISCED 5B>	☐5
f)	<ISCED 5A or 6>	☐6

Notes: For a definition of <**ISCED 2**> and <**ISCED 3**> see the note in Q7 of the student questionnaire.
For a definition of <**ISCED 4**> to <**ISCED 6**> levels see the note in Q14 of the student questionnaire.

Q18 (PA20) **What occupation do you expect your child to have when they are about 30 years old?**

Please write in the job title:

Parents' migration background

Q19 (PA21) **In what country were the following people in the child's family born?**
(Please tick one answer per column.)

	Mother	Father	Maternal Grand-mother	Maternal Grand-father	Paternal Grand-mother	Paternal Grand-father
<Test Country>	☐1	☐1	☐1	☐1	☐1	☐1
<Country A>	☐2	☐2	☐2	☐2	☐2	☐2
<Country B>	☐3	☐3	☐3	☐3	☐3	☐3
<Country C>	☐4	☐4	☐4	☐4	☐4	☐4
<Country D>	☐5	☐5	☐5	☐5	☐5	☐5
<Country E>	☐6	☐6	☐6	☐6	☐6	☐6
<Country F>	☐7	☐7	☐7	☐7	☐7	☐7

Notes: <**Test Country**> refers to the country being tested.
<**Country X**>, each country or economy may choose certain countries for this question.

Q20 (PA22)	**If the child's father was NOT born in <country of test>, how old was he when he arrived in <country of test>?** *(If less than 12 months old, please write zero (0).)*	
	Years	

Note: <**Country of test**> refers to the country being tested.

Q21 (PA23)	**If the child's mother was NOT born in <country of test>, how old was she when she arrived in <country of test>?** *(If less than 12 months old, please write zero (0).)*	
	Years	

Note: For a definition of <**Country of test**> see Q20.

Q22 (PA24) **In what country are the father and mother of the child legal citizens? If they are dual citizens, you may tick more than one country.**
(Please tick as many as apply.)

		Mother	Father
a)	<Test Country>	☐1	☐1
b)	<Country A>	☐2	☐2
c)	<Country B>	☐3	☐3
d)	<Country C>	☐4	☐4

Note: For a definition of <**Test country**> see Q19.

Q23 (PA25) **What language do the father and mother of the child speak at home most of the time?**
(Please tick one answer per column.)

	Mother	Father
<Test language>	☐1	☐1
<Other official national languages>	☐2	☐2
<Other national dialects or languages>	☐3	☐3
< Other language 1>	☐4	☐4
< Other language 2>	☐5	☐5
< Other language 3>	☐6	☐6
Other language	☐7	☐7

Note: <**Test language**> refers to the language of instruction in which you would administer the PISA reading assessment. It should be adapted to refer to the national name of the 'language of instruction' course or lessons.

Annex B

PISA 2012 EXPERT GROUPS

Annex B lists the members of the expert groups who were involved in developing the PISA 2012 framework for the major domain (mathematics), the new domains (problem solving and financial literacy) and the questionnaires. The lists of the experts involved in developing the PISA 2009 framework for reading and the 2006 framework for science can be found in the OECD publications *PISA 2009 Assessment Framework – Key competencies in Reading, Mathematics and Science* (2009) and *Assessing Scientific, Reading and Mathematical Literacy – A Framework for PISA 2006* (2006), respectively.

Mathematics expert group

Kaye Stacey - Chair
Melbourne Graduate School of Education
(Science and Mathematics Education)
University of Melbourne
Australia

Caroline Bardini
Melbourne Graduate School of Education
University of Melbourne
Australia

Werner Blum
Department of Mathematics
University of Kassel
Germany

Joan Ferrini-Mundy
College of Education
Michigan State University
United States of America

Solomon Garfunkel
COMAP
United States of America

Toshikazu Ikeda
Yokohama National University
Japan

Zbigniew Marciniak
Institute of Mathematics
Warsaw University
Poland

Mogens Niss
IMFUFA, Department of Sciences
Roskilde University
Denmark

Martin Ripley
World Class Arena Limited
United Kingdom

William Schmidt
College of Education
Michigan State University
United States of America

Problem solving expert group

Joachim Funke - Chair
Psychologisches Institut
University of Heidelberg
Germany

Beno Csapo
Institute of Education
University of Szeged
Hungary

John Dossey
Department of Mathematical Sciences
Illinois State University
United States of America

Arthur Graesser
Psychology, Computer Science, and Institute
for Intelligent Systems
The University of Memphis
United States of America

Detlev Leutner
Department of Instructional Psychology
Duisburg-Essen University
Germany

Romain Martin
Université de Luxembourg FLSHASE
Luxembourg

Richard Mayer
Department of Psychology
University of California, Santa Barbara
United States of America

Ming Ming Tan
Sciences, Curriculum Planning and Development Division
Ministry of Education
Singapore

Financial literacy expert group

Annamaria Lusardi – Chair
The George Washington University School of Business
United States of America

Jean-Pierre Boisivon
Université de Paris II Panthéon-Assas
France

Diana Crossan
Commission for Financial Literacy and Retirement Income
New Zealand

Peter Cuzner
Australian Securities and Investments Commission
Australia

Jeanne Hogarth
Federal Reserve System
United States of America

Dušan Hradil
Ministry of Finance
Czech Republic

Stan Jones
Consultant
Canada

Sue Lewis
Consultant
United Kingdom

Questionnaire expert group

Eckhard Klieme – Chair
Deutsches Institut für Internationale Pädagogische Forschung
(DIPF)
Frankfurt am Main, Germany

Eduardo Backhoff
University of Baja California at the Institute of Educational Research and Development
Mexico

Ying-yi Hong
Nanyang Business School of Nanyang Technological University
Singapore

David Kaplan
Department of Educational Psychology
University of Wisconsin
United States of America

Henry Levin
Teachers College, Columbia University
United States of America

Jaap Scheerens
Faculty of Educational, Science and Technology
University of Twente
Netherlands

William Schmidt
College of Education
Michigan State University
United States of America

Fons van de Vijver
Faculty of Social and Behavioural Studies
Tilburg University
Netherlands

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(98 2013 01 1P) ISBN 978-92-64-19052-8 – No. 60505 2013

PISA 2012 Assessment and Analytical Framework

MATHEMATICS, READING, SCIENCE, PROBLEM SOLVING AND FINANCIAL LITERACY

Are students well prepared to meet the challenges of the future? Can they analyse, reason and communicate their ideas effectively? Have they found the kinds of interests they can pursue throughout their lives as productive members of the economy and society? The OECD Programme for International Student Assessment (PISA) seeks to answer these questions through the most comprehensive and rigorous international assessment of student knowledge and skills. *PISA 2012 Assessment and Analytical Framework* presents the conceptual framework underlying the fifth cycle of PISA. Similar to the previous cycles, the 2012 assessment covers reading, mathematics and science, with the major focus on mathematical literacy. Two other domains are evaluated: problem solving and financial literacy. Students respond to a background questionnaire and, as an option, to an educational career questionnaire as well as another questionnaire about Information and Communication Technologies (ICTs). Additional supporting information is gathered from the school authorities through the school questionnaire and from the parents through a third optional questionnaire. Sixty-six countries and economies, including all 34 OECD member countries, are taking part in the PISA 2012 assessment.

Contents

THE OECD PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA)

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 merely with whether they have mastered specific curricular content. PISA's unique features include its:

- Policy orientation, which highlights differences in performance patterns and identifies features common to high-performing students, schools and education systems by linking data on learning outcomes with data on student characteristics and other key factors that shape learning in and outside of school.
- Innovative concept of "literacy", which refers both to students' capacity to apply knowledge and skills in key subject areas and to their ability to analyse, reason and communicate effectively as they pose, interpret and solve problems in a variety of situations.
- Relevance to lifelong learning, which goes beyond assessing students' competencies in school subjects by asking them to report on their 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 32 partner countries and economies.

Consult this publication on line at *http://dx.doi.org/10.1787/9789264190511-en*

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases.
Visit ***www.oecd-ilibrary.org*** for more information.

2013

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-19052-8
98 2013 01 1P