Programme for International Student Assessment

Problem Solving for Tomorrow's World

First Measures of Cross-Curricular Competencies from PISA 2003

OECD
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Pursuant to Article 1 of the Convention signed in Paris on 14th December 1960, and which came into force on 30th September 1961, the Organisation for Economic Co-operation and Development (OECD) shall promote policies designed:

– to achieve the highest sustainable economic growth and employment and a rising standard of living in member countries, while maintaining financial stability, and thus to contribute to the development of the world economy;
– to contribute to sound economic expansion in member as well as non-member countries in the process of economic development; and
– to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations.

The original member countries of the OECD are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The following countries became members subsequently through accession at the dates indicated hereafter: Japan (28th April 1964), Finland (28th January 1969), Australia (7th June 1971), New Zealand (29th May 1973), Mexico (18th May 1994), the Czech Republic (21st December 1995), Hungary (7th May 1996), Poland (22nd November 1996), Korea (12th December 1996) and the Slovak Republic (14th December 2000). The Commission of the European Communities takes part in the work of the OECD (Article 13 of the OECD Convention).

Publié en français sous le titre :

Résoudre des problèmes, un atout pour réussir – Premières évaluations des compétences transdisciplinaires issues de PISA 2003

Foreword

All stakeholders – parents, students, those who teach and run education systems as well as the general public – need to be informed on how well their education systems prepare students for life. Knowledge and skills in school subjects such as languages, mathematics and science are an essential foundation for this but a much wider range of competencies is needed for students to be well prepared for the future. Problem-solving skills, *i.e.* the capacity of students to understand problems situated in novel and cross-curricular settings, to identify relevant information or constraints, to represent possible alternatives or solution paths, to develop solution strategies, and to solve problems and communicate the solutions, are an example of this wider range of competencies.

The 2003 assessment of the Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA) included an assessment of students' problem-solving skills, providing for the first time a direct assessment of life competencies that apply across different areas of the school curriculum.

About one in five 15-year-olds in OECD countries can be considered a *reflective, communicative problem solver*. These students are able not only to analyse a situation and make decisions, they are also capable of managing multiple conditions simultaneously. They can think about the underlying relationships in a problem, solve it systematically, check their work and communicate the results. In some countries, more than a third of students reach this high level of problem-solving competencies. In other countries, however, the majority of students cannot even be classified as *basic problem solvers*, a level at which they are required to deal with only a single data source containing discrete, well-defined information.

How can countries raise their performance in this increasingly important competency area and what can countries with lower performance levels learn from those where students do well? This report seeks to answer such questions. It complements *Learning for Tomorrow's World – First Results from PISA 2003,* which focuses on knowledge and skills in mathematics, science and reading, and it goes beyond an examination of the relative standing of countries in students' problem-solving performance by considering how problem-solving performance relates to learning outcomes in other areas and how it varies between the genders and between socio-economic groups. It also provides insights into some of the factors that are associated with the development of problem-solving skills and into how these factors interact and what the implications are for policy development. Most importantly, the report sheds light on countries that succeed

in achieving high performance standards while at the same time providing an equitable distribution of learning opportunities. Results in these countries pose challenges for other countries by showing what it is possible to achieve.

The report is the product of a collaborative effort between the countries participating in PISA, the experts and institutions working within the framework of the PISA Consortium, and the OECD. The report was drafted by John Dossey, Johannes Hartig, Eckhard Klieme and Margaret Wu, under the direction of the OECD Directorate for Education, principally by Claire Shewbridge and Andreas Schleicher, with advice and analytic support from Raymond Adams, Barry McCrae and Ross Turner. The PISA problem-solving framework and assessment instruments were prepared by the PISA Consortium and PISA Problem Solving Expert Group under the direction of Raymond Adams at the Australian Council for Educational Research. Data analytic support was provided by Alla Berezener, Johannes Hartig and Margaret Wu.

The development of the report was steered by the PISA Governing Board, which is chaired by Ryo Watanabe (Japan). Annex C of the report lists the members of the various PISA bodies as well as the individual experts and consultants who have contributed to this report and to PISA in general.

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

Ryo Watanabe
Chair of the PISA Governing Board

Barry McGaw
Director for Education, OECD

Table of Contents

ANNEX B

ANNEX C

LIST OF BOXES

LIST OF FIGURES

LIST OF TABLES

1

PISA 2003 and Problem Solving

Introduction

This report looks at how well students can solve problems not linked to specific parts of the school curriculum.

How well prepared are young adults to solve the problems that they will encounter in life beyond school, in order to fulfil their goals in work, as citizens and in further learning? For some of life's challenges, they will need to draw on knowledge and skills learned in particular parts of the school curriculum – for example, to recognise and solve a mathematics-related problem. Other problems will be less obviously linked to school knowledge, and will often require students to deal with unfamiliar situations by thinking flexibly and creatively. This report is concerned with problem solving of the second, more general variety.

It should be understood both as a part of the initial results of PISA 2003…

The Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA) conducted its second survey of student knowledge and skills of 15-year-olds in 2003. *Learning for Tomorrow's World – First Results from PISA 2003* (OECD, 2004a) summarises the results from the assessment of mathematics, science and reading. This report summarises results from the assessment of the problem-solving skills. This feature of PISA represents an important development in an innovative international survey seeking to probe beyond conventional assessments of student abilities centred on particular school subject areas.

…and in relation to PISA as a whole.

PISA's assessment of problem-solving skills needs to be understood in the context of the overall features and purposes of PISA. The introduction to *Learning for Tomorrow's World – First Results from PISA 2003* (OECD, 2004a) describes the survey and explains how PISA assesses mathematics, science and reading. A brief summary of key features of PISA is provided below before this report turns to how PISA assesses problem-solving skills.

PISA measures how well 15-year-olds are prepared for life's challenges.

PISA seeks to measure how well young adults, at age 15 – and therefore approaching the end of compulsory schooling – are prepared to meet the challenges of today's knowledge societies. The assessment is forward-looking, focusing on young people's ability to use their knowledge and skills to meet real-life challenges, rather than just examining the extent to which they have mastered a specific school curriculum. This orientation reflects a change in the goals and objectives of curricula themselves, which are increasingly concerned with how students use what they learn at school, and not merely whether they can reproduce what they have learned. Key features driving the development of PISA have been:

- its policy orientation, with design and reporting methods determined by the need of governments to draw policy lessons;
- the innovative "literacy" concept that 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;

- its 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 motivation to learn, their beliefs about themselves and their learning strategies;
- its regularity, which will enable countries to monitor their progress in meeting key learning objectives; and
- its breadth of geographical coverage, with the 48 countries that have participated in a PISA assessment so far and the 11 additional ones that will join the PISA 2006 assessment, representing a total of one-third of the world population and almost nine-tenths of the world's GDP.[1]

Helped by leading experts, participating countries and the OECD have created valid cross-country assessment materials.

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 steered jointly by their 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 most developed countries, as well as in many others at earlier stages of economic development.

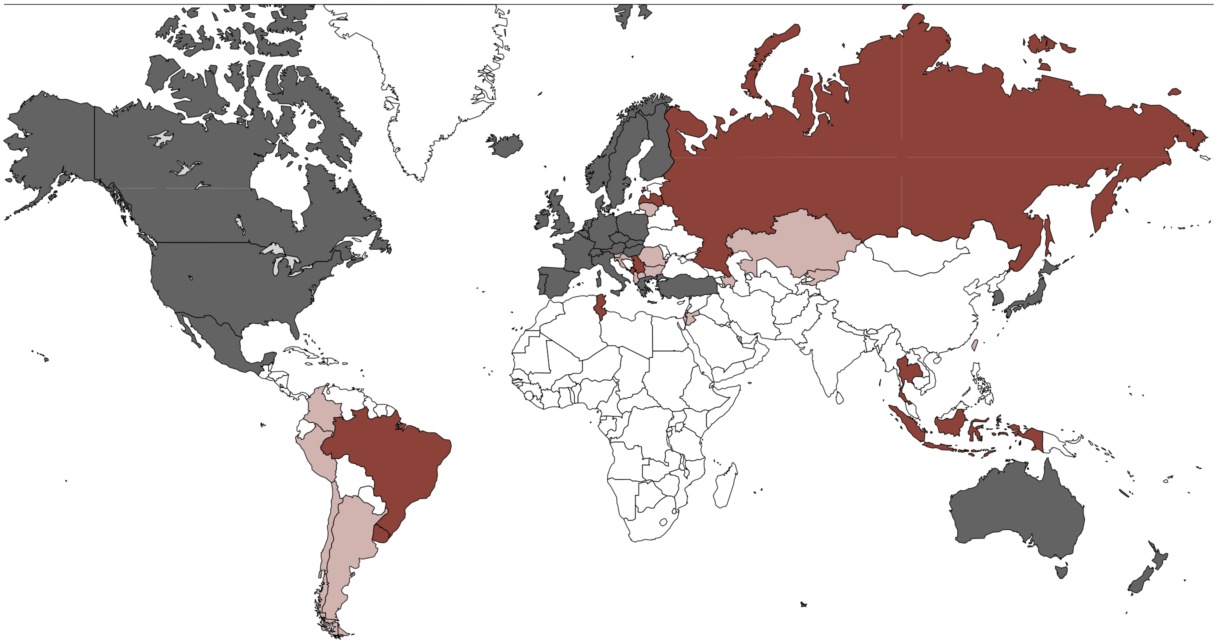
The first survey took place in 2000 and focused on reading literacy, while PISA 2003 focused on mathematics and PISA 2006 will focus on science.

The first PISA survey was conducted in 2000 in 32 countries (including 28 OECD member countries) and repeated in 11 additional partner countries in 2002. In PISA 2000, where the focus was on reading, students performed written tasks under independently supervised test conditions in their schools. The first results were published in 2001 (OECD, 2001a) and 2003 (OECD, 2003a), and followed by a series of thematic reports looking in more depth at various aspects of the results.[2] PISA 2003, reported on here, was conducted in 41 countries, including all 30 OECD member countries (Figure 1.1). It included an in-depth assessment of mathematics as well as less detailed assessments in science and reading. A special feature of the 2003 survey was the one-off assessment of problem-solving skills. In the next three-yearly survey, PISA 2006, the primary focus will be on science, and there will be a return to the focus on reading in 2009.[3]

PISA is being used not just in the OECD area but across the world.

Although PISA was originally created by the OECD governments in response to their own needs, it has now become a major policy tool for many other countries and economies as well. PISA is playing an increasing role in regions around the world, and the survey has now been conducted or is planned in the partner countries in Southeast Asia (Hong Kong-China, Indonesia, Macao-China, Chinese Taipei and Thailand), Eastern Europe (Albania, Bulgaria, Croatia, Estonia, Latvia,

Figure 1.1 ■ **A map of PISA countries**



OECD countries

Australia
Austria
Belgium
Canada
Czech Republic
Denmark
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Japan
Korea
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States

Partner countries in PISA 2003

Brazil
Hong Kong-China
Indonesia
Latvia
Liechtenstein
Macao-China
Russian Federation
Serbia and Montenegro
Thailand
Tunisia
Uruguay

Partner countries in other PISA assesments

Albania
Argentina
Azerbaijan
Bulgaria
Chile
Colombia
Croatia
Estonia
Israel
Jordan
Kazakhstan
Kyrgyz Republic
Lithuania
Macedonia
Peru
Qatar
Romania
Slovenia
Chinese Taipei

Lithuania, the Former Yugoslav Republic of Macedonia, Romania, the Russian Federation, Serbia[4] and Slovenia), the Middle East (Jordan, Israel and Qatar), South America (Argentina, Brazil, Chile, Colombia, Peru and Uruguay) and North Africa (Tunisia). Across the world, policy makers use PISA findings to:

- gauge the literacy 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 system.

Globally, it has become part of the public debate.

National interest in PISA is illustrated by the many reports produced in participating countries and by the numerous references to the results of PISA in public debates and the media throughout the world (see *www.pisa.oecd.org* for examples).

Box 1.1 ■ **Key features of the PISA 2003 assessment**

Content

- The survey covers mathematics (the main focus in 2003), reading, science and problem solving. PISA considers student knowledge in these areas not in isolation but in relation to students' ability to reflect on their knowledge and experience and to apply them to real world issues. The emphasis is on the mastery of processes, the understanding of concepts, and the ability to function in various situations within each assessment area.
- PISA integrates the assessment of subject-specific knowledge with cross-curricular competencies. In PISA 2003, as in 2000, students assessed their own characteristics as learners. The 2003 survey also introduced the first assessment of wider student competencies – assessing problem-solving abilities.

Methods

- Each participating student spent two hours carrying out pencil-and-paper tasks.
- Questions requiring students to construct their own answers were combined with multiple-choice items. Items were typically organised in units based on a written passage or graphic, of the kind that students might encounter in real life.
- A total of six-and-a-half hours of assessment items was included, with different students taking different combinations of the assessment items. Three-and-a-half hours of testing time was in mathematics, with one hour each for reading, science and problem solving.
- Students answered a questionnaire that took about 30 minutes to complete and focused on their background, their learning habits and their perceptions of the learning environment, as well as on their engagement and motivation.
- School principals completed a questionnaire about their school that included demographic characteristics as well as an assessment of the quality of the learning environment at school.

Outcomes

- A profile of knowledge and skills among 15-year-olds in 2003.
- Contextual indicators relating performance results to student and school characteristics.
- A knowledge base for policy analysis and research.
- A first estimate of change in student knowledge and skills over time, between the assessments in 2000 and 2003.

Sample size

- Well over a quarter of a million students, representing about 23 million 15-year-olds in the schools of the 41 participating countries, were assessed on the basis of scientific probability samples.

Future assessments

- The PISA 2006 assessment will focus on science and PISA 2009 will return to a focus on reading.
- Part of future assessments will require students to use computers, expanding the scope of the skills that can be tested and reflecting the importance of information and computer technology (ICT) as a medium in modern societies.

Problem solving in PISA 2003

A framework has been established to enable countries to assess students' ability to solve problems that are not bound to specific areas of school knowledge.

The collection of data concerning students' problem-solving skills as part of PISA 2003 was undertaken because the OECD countries attach great importance to how far students' capabilities in reading, mathematics and science are matched by an overall capability to solve problems in real-life situations beyond the specific context of school subject areas. To address this, the OECD countries established a framework and assessment instruments to evaluate students' capacities to:

- identify problems in cross-curricular settings;
- identify relevant information or constraints;
- represent possible alternatives or solution paths;
- select solution strategies;
- solve problems;
- check or reflect on the solutions; and
- communicate the results.

The framework for this assessment is discussed in Chapter 2 and described in full in *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003b).

PISA chose three types of problem-solving exercises to assess.

Given the amount of time available for the assessment, the decision was made to focus on students' problem-solving capabilities in three types of situation:

- making decisions under constraints;
- evaluating and designing systems for a particular situation; and
- trouble-shooting a malfunctioning device or system based on a set of symptoms.

Working with these types of problems, a large set of tasks was developed and field tested in participating countries. The results were 19 tasks that required problem-solving skills, most of which are set in units consisting of two or three related items dealing with the same contextual situation. For example, the unit *Holiday* (shown below) consists of two items – the first asking students a direct question that assesses to what degree they understand the problem and are able to grasp the scheduling decisions that must be made, the second question asking for an itinerary that meets the criteria given. In responding, students have to deal with the constraints of the roads, distances, camp locations, towns that the individual (Zoe) wants to visit; the maximum amounts of travel per day; and the visiting times in the specific towns she wants to visit on her trip.

HOLIDAY

This problem is about planning the best route for a holiday.

Figures 1 and 2 show a map of the area and the distances between towns.

Figure 1. Map of roads between towns

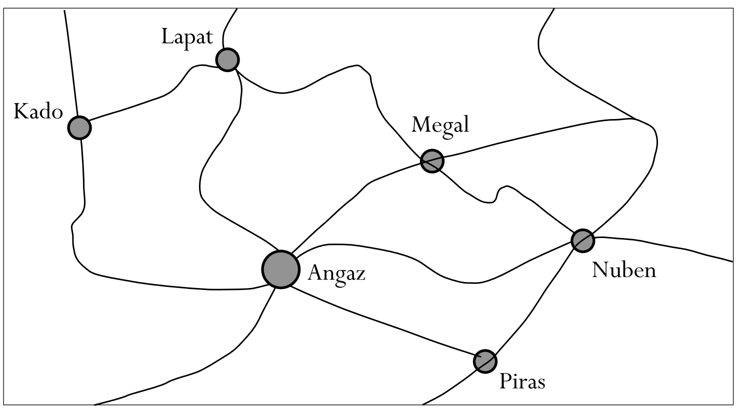


Figure 2. Shortest road distance of towns from each other in kilometres.

	Angaz	Kado	Lapat	Megal	Nuben	Piras
Angaz						
Kado	550					
Lapat	500	300				
Megal	300	850	550			
Nuben	500		1000	450		
Piras	300	850	800	600	250	

HOLIDAY – Question 1

Calculate the shortest distance by road between Nuben and Kado.

Distance: .... kilometres.

HOLIDAY – Question 2

Zoe lives in Angaz. She wants to visit Kado and Lapat. She can only travel **up to 300 kilometres** in any one day, but can break her journey by camping overnight anywhere between towns.

Zoe will stay for **two nights** in each town, so that she can spend one whole day sightseeing in each town.

Show Zoe's itinerary by completing the following table to indicate where she stays each night.

Day	Overnight Stay
1	Camp-site between Angaz and Kado.
2	
3	
4	
5	
6	
7	Angaz

These are described in more detail in Chapter 4.

All of the items in the units for problem solving are shown in Chapter 4, along with the criteria used to evaluate student performance. Each of the items is illustrated along with a sample of student work, and the difficulty of each item is matched with a score on a scale constructed to report problem-solving performance among students participating in PISA 2003.

The information on problem solving enriches our understanding of student competencies…

The data from this part of the PISA assessment give a first glimpse of what students can do when asked to use their total accumulated knowledge and skills to solve problems in authentic situations that are not associated with a single part of the school curriculum.

…and can be used in combination with other PISA results to inform the development of school systems.

The results from PISA provide a basis for the participating countries to compare the results of their varied investments in education and learning. When diverse educational structures are compared in terms of their student outcomes, some patterns of similarity emerge. Analyses of the outcomes suggest possible alternatives for action within the countries or support for continued work along the path that has been chosen for education within the countries. Most importantly, the findings provide those responsible for education with information through which they can examine the strengths and weaknesses of the programmes they are currently offering their students.

PISA assesses students aged 15 who are still at school, regardless of grade or institution…

In order to ensure the comparability of the results across countries, PISA needs to assess comparable target populations. Differences between countries in the nature and extent of pre-primary education and care, in the age of entry to formal schooling, and in the structure of the education system do not allow

school grades to be defined so that they are internationally comparable. Valid international comparisons of educational performance must, therefore, define their populations with reference to a target age. PISA covers students who are aged between 15 years 3 months and 16 years 2 months at the time of the assessment, regardless of the grade or type of institution in which they are enrolled and of whether they are in full-time or part-time education. The use of this age in PISA, across countries and over time, allows a consistent comparison of the performance of students shortly before they complete compulsory education.

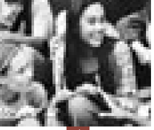
As a result, this report is able to make statements about the knowledge and skills of individuals born in the same year and still at school at 15 years of age, but having differing educational experiences, both within and outside school. The number of school grades in which these students are to be found depends on a country's policies on school entry and promotion. Furthermore, in some countries, students in the PISA target population represent different education systems, tracks or streams.

…and only leaves out small parts of the target population…

Stringent technical standards were established for the definition of national target populations. PISA excludes 15-year-olds not enrolled in educational institutions. In the remainder of this report "15-year-olds" is used as shorthand to denote the PISA student population. Coverage of the target population of 15-year-olds within education is very high compared with other international surveys: relatively few schools were ineligible for participation, for example because of geographically remoteness or because their students had special needs. In 24 out of 41 participating countries, the percentage of school-level exclusions amounted to less than 1 per cent, and to less than 3 per cent in all countries except Mexico (3.6 per cent), Switzerland (3.4 per cent), the United Kingdom (3.4 per cent) and the partner countries Latvia (3.8 per cent) and Serbia (5.3 per cent). When accounting for the exclusion within schools of students who met certain internationally established criteria,[5] the exclusion rates increase slightly. However, it remains below 2 per cent in 19 participating countries, below 4 per cent in 29 participating countries, below 6 per cent in all but two countries and below 8 per cent in all countries (Annex A3). This high level of coverage contributes to the comparability of the assessment results. For example, even assuming that the excluded students would have systematically scored worse than those who participated, and that this relationship is moderately strong, an exclusion rate in the order of 5 per cent would likely lead to an overestimation of national mean scores of less than 5 score points.[6] Moreover, in most cases the exclusions were inevitable. For example, in New Zealand 2.3 per cent of the students were excluded because they had less than one year of instruction in English (often because they were foreign fee-paying students) and were therefore not able to follow the instructions of the assessment.

…with sufficiently large scientific samples to allow for valid comparisons.

The specific sample design and size for each country was designed to maximise sampling efficiency for student-level estimates. In OECD countries, sample sizes ranged from 3 350 students in Iceland to 30 000 students in Mexico.



This selection of samples was monitored internationally and accompanied by rigorous standards for the participation rate to ensure that the PISA results reflect the skills of 15-year-old students in participating countries.

Organisation of this report

This report describes and analyses student performance in problem solving.

The report provides an in-depth examination of the results on the performance of students in the 41 countries participating in PISA 2003 on the items for problem solving. The following four chapters provide detailed analysis of the data, their meaning and their implications.

Chapter 2 describes the criteria used to assess it, and reports overall country performance.

Chapter 2 provides an introduction to problem solving and a closer inspection of the definition of the assessment area as used by PISA 2003 in the development of the assessment. Central to this description is the role that problem solving plays as a basis for future learning, for fruitful employment, and for productive citizenship. Following a further description of the assessment framework through a selection of sample problems, the PISA problem-solving scale is discussed using student performance on these problems as a way of interpreting the scale. This is followed by an overall discussion of the performance of students from the 41 participating nations.

Chapter 3 compares student performance in problem solving to their performance in other PISA assessment areas.

Chapter 3 analyses students' results in problem solving, mathematics, reading and science to better understand the cognitive demands of the problem-solving assessment. The chapter provides a country-by-country comparison of mean performance of students and compares this with their mean performances in mathematics, reading and science.

Chapter 4 looks in more detail at how students responded to individual items.

Chapter 4 provides a comprehensive look at the problem-solving assessment. It describes the tasks and individual items classified by PISA problem types. Several items are accompanied by sample student work illustrating the criteria for scoring and the variety of problem-solving approaches that students used in their solutions.

Chapter 5 analyses how student competencies in problem solving relate to gender and family background.

Chapter 5 provides an analysis of the relationships between problem-solving performance and a variety of student, family, and other background characteristics. Central to these comparisons is the consideration of gender differences in problem solving. This is followed by consideration of the impact of student family features on student problem solving. These analyses include the occupational status of students' parents and other factors having central importance to students' performance on the problem-solving items.

Notes

1. The combined population of all countries (Chinese Taipei not included) that participate in the PISA 2000, 2003 or 2006 assessments amounts to 32 per cent of the 2002 world population. The combined GDP of these countries amounts to 87.4 per cent of the 2002 world GDP. The data on GDP and population sizes were derived from the U.N. World Development Indicators database.
2. Themes of international thematic reports included: *Reading for Change – Performance and Engagement Across Countries* (OECD, 2002a), *Learners for Life – Student Approaches to Learning* (OECD, 2003c), *Student Engagement at School – A Sense of Belonging and Participation* (OECD, 2003d), *What Makes School Systems Perform* (OECD, 2004b) and *School Factors Relating to Quality and Equity* (OECD, forthcoming)
3. The framework for the PISA 2006 assessment has been finalised and preparations for the implementation of the assessment are currently underway. Governments will decide on subsequent PISA assessments in 2005.
4. For the country Serbia and Montenegro, data for Montenegro are not available. The latter accounts for 7.9 per cent of the national population. The name "Serbia" is used as a shorthand for the Serbian part of Serbia and Montenegro.
5. Countries were permitted to exclude up to 2.5 per cent of the national desired target population within schools if these students were: *i)* considered in the professional opinion of the school principal or of other qualified staff members, to be educable mentally retarded or who had been defined as such through psychological tests (including students who were emotionally or mentally unable to follow the general instructions given in PISA); *ii)* permanently and physically disabled in such a way that they could not perform in the PISA assessment situation (functionally disabled students who could respond were to be included in the assessment); or *iii)* non-native language speakers with less than one year of instruction in the language of the assessment (for details see Annex A3).
6. If the correlation between the propensity of exclusions and student performance is 0.3, resulting mean scores would likely be overestimated by 1 score point if the exclusion rate is 1 per cent, by 3 score points if the exclusion rate is 5 per cent, and by 6 score points if the exclusion rate is 10 per cent. If the correlation between the propensity of exclusions and student performance is 0.5, resulting mean scores would be overestimated by 1 score point if the exclusion rate is 1 per cent, by 5 score points if the exclusion rate is 5 per cent, and by 10 score points if the exclusion rate is 10 per cent. For this calculation, a model was employed that assumes a bivariate normal distribution for the propensity to participate and performance. For details see the *PISA 2000 Technical Report* (OECD 2002b).

READERS' GUIDE

Data underlying the figures

The data referred to in Chapters 2, 3 and 5 of this report are presented in Annex B and, with additional detail, on the web site *www.pisa.oecd.org*. Three symbols are used to denote missing data:

a The category does not apply in the country concerned. Data are therefore missing.

c There are too few observations to provide reliable estimates (*i.e.* there are fewer than 3 per cent of students for this cell or too few schools for valid inferences). However, these statistics were included in the calculation of cross-country averages.

m Data are not available. These data were collected but subsequently removed from the publication for technical reasons.

Calculation of international averages

An OECD average was calculated for most indicators presented in this report. In the case of some indicators, a total representing the OECD area as a whole was also calculated:

- The ***OECD average*** takes the OECD countries as a single entity, to which each country contributes with equal weight. For statistics such as percentages of mean scores, the OECD average corresponds to the arithmetic mean of the respective country statistics. In contrast, for statistics relating to variation, the OECD average may differ from the arithmetic mean of the country statistics because it not only reflects variation within countries, but also variation that lies between countries.
- The ***OECD total*** takes the OECD countries as a single entity, to which each country contributes in proportion to the number of 15-year-olds enrolled in its schools (see Annex A3 for data). It illustrates how a country compares with the OECD area as a whole.

In this publication, the OECD total is generally used when references are made to the stock of human capital in the OECD area. Where the focus is on comparing performance across education systems, the OECD average is used. In the case of some countries, data may not be available for specific indicators or specific categories may not apply. Readers should, therefore, keep in mind that the terms ***OECD average*** and ***OECD total*** refer to the OECD countries included in the respective comparisons. All international averages include data for the United Kingdom, even where these data, for reasons explained in Annex A3, are not shown in the respective data tables.

Rounding of figures

Because of rounding, some figures in tables may not exactly add up to the totals. Totals, differences and averages are always calculated on the basis of exact numbers and are rounded only after calculation.

When standard errors in this publication have been rounded to one or two decimal places and the value 0.0 or 0.00 is shown, this does not imply that the standard error is zero, but that it is smaller than 0.05 or 0.005 respectively.

Reporting of student data

The report usually uses "15-year-olds" as shorthand for the PISA target population. In practice, this refers to students who were aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months at the beginning of the assessment period and who were enrolled in an educational institution, regardless of the grade level or type of institution, and of whether they were attending full-time or part-time (for details see Annex A3).

Abbreviations used in this report

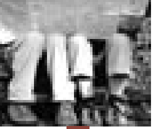
The following abbreviations are used in this report:

GDP Gross Domestic Product
ISCED International Standard Classification of Education
SD Standard deviation
SE Standard error

Further documentation

For further information on the PISA assessment instruments and the methods used in PISA, see the *PISA 2000 Technical Report* (OECD, 2002b) and the PISA Web site (*www.pisa.oecd.org*).

Problem Solving in PISA 2003 – How It Was Measured and How Students Performed



Introduction

This chapter describes how PISA measured problem solving and summarises student performance overall.

This chapter provides an overview of how students' performance in problem solving was measured in PISA 2003, reports on how many students reached various levels of proficiency and gives the mean and distribution of performance in each participating country.

- First, the chapter defines problem solving, reviews the kind of problem-solving tasks that were used in PISA 2003 and describes the requirements made of students in solving these problems.
- Second, the chapter describes the way in which student performance in problem solving was measured. This is illustrated in relation to items used in this assessment, and the percentage of each country's students at each proficiency level of the problem-solving scale is reported.
- Third, the chapter summarises the performance of students in each of the countries participating in PISA 2003 by reporting their mean performance and describing the distribution of scores on the problem-solving assessment for the students within each country.

Problem solving in PISA

Problem solving is a central part of education across the curriculum.

Curricula in various subject areas often call for students to confront problem situations by understanding information that is given, identifying critical features and any relationships in a situation, constructing or applying one or more external representations, resolving ensuing questions and, finally, evaluating, justifying and communicating results as a means to further understanding the situation. This is because problem solving is widely seen as providing an essential basis for future learning, for effectively participating in society, and for conducting personal activities.

To assess it requires tasks that are…

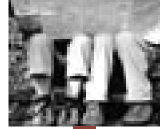
The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills (OECD, 2003b) through which OECD countries established the guiding principles for comparing problem-solving performance across countries in PISA, defines problem competencies as:

> … 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 content areas or curricular areas that might be applicable are not within a single subject area of mathematics, science or reading.

Several aspects of this definition are worth noting.

…situated in real-life contexts…

- The first is that the settings for the problems should be real. They should draw on situations that represent contexts that could conceivably occur in a student's life or, at least, be situations the student can identify as being important to society, if not directly applicable to his or her personal life. Thus, a real-life problem calls on individuals to merge knowledge and strategies to confront and resolve a problem, when the method by which this needs to be accomplished is not readily apparent to the problem solver.



- The second feature is that they are not immediately resolvable through the application of some defined process that the student has studied, and probably practised, at school. The problems should present new types of questions requiring the student to work out what to do. This is what causes the item really to be a problem-solving item. Such problems call on individuals to move among different, but sometimes related, representations and to exhibit a certain degree of flexibility in the ways in which they access, manage, evaluate and reflect on information.

…not resolvable through the application of routine solutions…

- Finally, the problems used should not be limited to a single content area that students would have studied and practised as part of their study of a single school subject in school.

…and require connections between multiple content areas.

Organisation of the assessment area

With this definition of problem solving, the nature of the tasks to be used in the assessment was established in *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003b), based on the following components:

The problem-solving tasks were defined by the …

- *Problem types.* PISA 2003 focused on three problem types: *decision making*, *system analysis and design,* and *trouble shooting*. These were chosen because they are widely applicable and occur in a variety of settings. The problem types used for PISA are described in more detail in the next section.

…the type of problem …

- *Problem context.* The problems used in the assessment were not set in the classroom or based on materials studied in the curriculum, but rather set in contexts that a student would find in his/her personal life, work and leisure, and in the community and society.

…the problem context…

- *Problem-solving processes*. The assessment was designed such that the results would describe the degree to which students are able to confront, structure, represent and solve problems effectively. Accordingly, the tasks included in the assessment were selected to collect evidence of students' knowledge and skills associated with the problem-solving process. In particular, students had to demonstrate that they could:

…and the problem-solving processes involved.

  - *Understand the problem:* This included understanding text, diagrams, formulas or tabular information and drawing inferences from them; relating information from various sources; demonstrating understanding of relevant concepts; and using information from students' background knowledge to understand the information given.

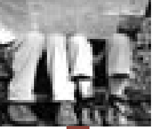
Students had to show their ability to understand the problem…

  - *Characterise the problem:* This included identifying the variables in the problem and noting their interrelationships; making decisions about which variables are relevant and irrelevant; constructing hypotheses; and retrieving, organising, considering and critically evaluating contextual information.

…identify the variables involved and their interrelationships…

  - *Represent the problem:* This included constructing tabular, graphical, symbolic or verbal representations; applying a given external representation to the solution of the problem; and shifting between representational formats.

…represent the problem…



…solve the problem…

– *Solve the problem:* This included making decisions (in the case of *decision making*); analysing a system or designing a system to meet certain goals (in the case of *system analysis and design*); and diagnosing and proposing a solution (in the case of *trouble shooting*).

…reflect on the solution…

– *Reflect on the solution:* This included examining solutions and looking for additional information or clarification; evaluating solutions from different perspectives in an attempt to restructure the solutions and making them more socially or technically acceptable; and justifying solutions.

…and communicate it.

– *Communicate the problem solution:* This included selecting appropriate media and representations to express and to communicate solutions to an outside audience.

Beyond drawing on a student's knowledge, good problems also draw upon their reasoning skills.

Beyond drawing on a student's knowledge, good problems also draw upon their 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. In solving a problem and communicating the results, the problem solver may need to organise information in a logical manner. These activities often require analytical reasoning, quantitative reasoning, analogical reasoning and combinatorial reasoning skills.

Thus, a student needs to combine many different cognitive processes to solve a problem and the PISA problem-solving assessment strives to identify the processes students use and to describe and quantify the quality of the students' work in problem solving.

Problems chosen for the PISA problem-solving assessment

Problem types included tasks related to decision making, system analysis and design and trouble shooting.

Three types of problem were chosen for the PISA problem-solving assessment: *decision making*, *system analysis and design* and *trouble shooting.* Figure 2.1 compares the features of each problem type. The three features outlined in the table (goals, processes and sources of complexity) serve as the basis for establishing a scale to describe increasing student proficiency in problem solving. The PISA problem-solving scale provides a representation of students' capacity to understand, characterise, represent, solve, reflect on and communicate their solutions to a problem.

The PISA problem-solving scale

Student performance was rated on a scale based on aspects of the above framework, with three levels of proficiency distinguishing between…

The PISA problem-solving scale derives from an analysis of the theoretical constructs underlying the problem-solving components detailed in Figure 2.1 and was validated by an analysis of student work on related tasks. The scale runs from students with the weakest problem-solving skills to those with the strongest problem-solving skills and has three distinct, described performance levels. These are referred to as proficiency levels, and provide an analytical model for describing what individual students are capable of, as well as comparing and contrasting student proficiency across countries.

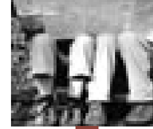


Figure 2.1 ■ **Features of the three types of problem solving**

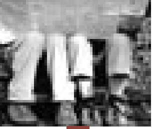
	Decision making	System analysis and design	Trouble shooting
Goals	Choosing among alternatives under constraints	Identifying the relationships between parts of a system and/or designing a system to express the relationships between parts	Diagnosing and correcting a faulty or underperforming system or mechanism
Processes involved	Understanding a situation where there are several alternatives and constraints and a specified task	Understanding the information that characterises a given system and the requirements associated with a specified task	Understanding the main features of a system or mechanism and its malfunctioning, and the demands of a specific task
	Identifying relevant constraints	Identifying relevant parts of the system	Identifying causally related variables
	Representing the possible alternatives	Representing the relationships among parts of the system	Representing the functioning of the system
	Making a decision among alternatives	Analysing or designing a system that captures the relationships between parts	Diagnosing the malfunctioning of the system and/or proposing a solution
	Checking and evaluating the decision	Checking and evaluating the analysis or the design of the system	Checking and evaluating the diagnosis/solution
	Communicating or justifying the decision	Communicating the analysis or justifying the proposed design	Communicating or justifying the diagnosis and the solution
Possible sources of complexity	Number of constraints	Number of interrelated variables and nature of relationships	Number of interrelated parts of the system or mechanism and the ways in which these parts interact
	Number and type of representations used (verbal, pictorial, numerical)	Number and type of representations used (verbal, pictorial, numerical)	Number and type of representations used (verbal, pictorial, numerical)

Level 3: Reflective, communicative problem solvers

Students proficient at Level 3 score above 592 points on the PISA problem-solving scale and typically do not only analyse a situation and make decisions, but also think about the underlying relationships in a problem and relate these to the solution. Students at Level 3 approach problems systematically, construct their own representations to help them solve it and verify that their solution satisfies all requirements of the problem. These students communicate their solutions to others using accurate written statements and other representations.

...reflective problem solvers that do not only analyse a situation and make correct decisions but also think about underlying relationships and relate these to solutions...

Students at Level 3 tend to consider and deal with a large number of conditions, such as monitoring variables, accounting for temporal restrictions, and other constraints. Problems at this level are demanding and require students to regulate their work. Students at the top of Level 3 can cope with multiple interrelated conditions that require students to work back and forth between their solution and the conditions laid out in the problem. Students at this level



organise and monitor their thinking while working out their solution. Level 3 problems are often multi-faceted and require students to manage all interactions simultaneously and develop a unique solution, and students at Level 3 are able to address such problems successfully and communicate their solutions clearly.

Students at Level 3 are also expected to be able to successfully complete tasks located at lower levels of the PISA problem-solving scale.

Level 2: Reasoning, decision-making problem solvers

…reasoning, decision-making problem-solvers…

Students proficient at Level 2 score from 499 to 592 points on the problem-solving scale and use reasoning and analytic processes and solve problems requiring decision-making skills. These students can apply various types of reasoning (inductive and deductive reasoning, reasoning about causes and effects, or reasoning with many combinations, which involves systematically comparing all possible variations in well-described situations) to analyse situations and to solve problems that require them to make a decision among well-defined alternatives. To analyse a system or make decisions, students at Level 2 combine and synthesise information from a variety of sources. They are able to combine various forms of representations (*e.g.* a formalised language, numerical information, and graphical information), handle unfamiliar representations (*e.g.* statements in a programming language or flow diagrams related to a mechanical or structural arrangement of components) and draw inferences based on two or more sources of information.

Students at Level 2 are also expected to be able to successfully complete tasks located at Level 1 of the PISA problem-solving scale.

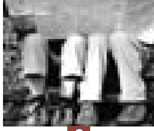
Level 1: Basic problem solvers

…and basic problem solvers.

Students proficient at Level 1 score from 405 to 499 points on the problem-solving scale and typically solve problems where they have to deal with only a single data source containing discrete, well-defined information. They understand the nature of a problem and consistently locate and retrieve information related to the major features of the problem. Students at Level 1 are able to transform the information in the problem to present the problem differently, *e.g.* take information from a table to create a drawing or graph. Also, students can apply information to check a limited number of well-defined conditions within the problem. However, students at Level 1 do not typically deal successfully with multi-faceted problems involving more than one data source or requiring them to reason with the information provided.

Below Level 1: Weak or emergent problem solvers

The PISA problem-solving assessment was not designed to assess elementary problem-solving processes. As such, the assessment materials did not contain sufficient tasks to describe fully performances that fall below Level 1. Students with performances below Level 1 have scores of less than 405 points on the problem-solving scale and consistently fail to understand even the easiest items in the assessment or fail to apply the necessary processes to characterise important features or represent the problems. At most, they can deal with straightforward problems with carefully structured tasks that require the students to give

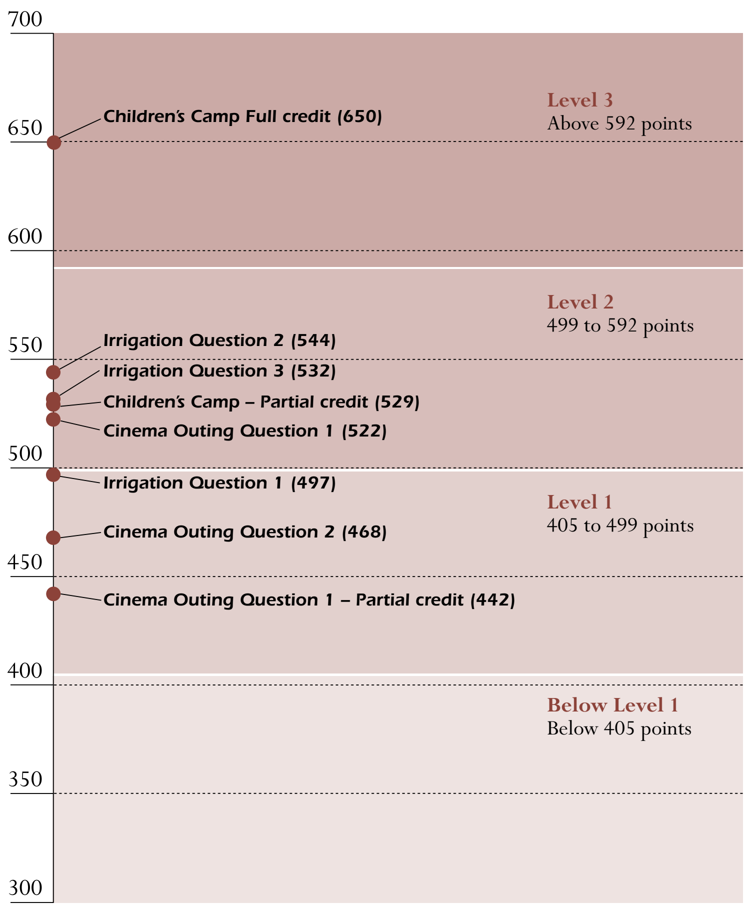


responses based on facts or to make observations with few or no inferences. Students below Level 1 have significant difficulties in making decisions, analysing or evaluating systems, and trouble-shooting situations.

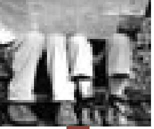
The three levels of problem solving are associated with a defined range of scores on the PISA problem-solving scale. In Figure 2.2 this scale is represented as a vertical line, with students' scores representing their level of problem-solving proficiency. A student can score full, partial or no credit for a given item. Scores for full or partial credit (including two levels of partial credit on one of the items) are expressed in terms of particular scores along the scale. Each assessment item is assigned a score, such that the majority of students with this score could expect to get the item correct. The mean student performance across OECD countries, weighted equally, was set at 500 score points, and the standard deviation was set at 100 score points. Thus, approximately two-thirds of student performances fall between 400 and 600 score points.

These proficiency levels are represented on a scale for which the mean score is 500 points and two-thirds score between 400 and 600 points.

Figure 2.2 ■ **The PISA problem-solving scale**



The three items shown below illustrate the nature of the various problem types and the processes required for students to succeed in problem-solving tasks at various levels of difficulty.



Decision making – the **CINEMA OUTING** problem

Context: Personal

Levels: Level 1 (*Cinema Outing,* Question 2) and
Level 2 (*Cinema Outing,* Question 1)

PISA scale score: 468 (*Cinema Outing,* Question 2) and
522 (*Cinema Outing,* Question 1)

Cinema Outing is a *decision-making* problem that presents students with a significant amount of information and a set of well-defined decisions to make based on the information given. Students proficient at Level 2 will typically be able to respond correctly to *Cinema Outing,* Question 1. Such students are capable of making decisions while considering a wide variety of boundary constraints and reasoning through what works and what does not work. Most of the decisions require the use of two or more pieces of the provided information. In addition, the student has to merge information from boundary conditions in the stated context, *e.g.* information about individuals' weekly schedules, commitments, and movies they had already seen, as well as noting which movies are showing, the showing times and film lengths, and the film ratings. *Cinema Outing,* Question 2 is a less demanding task. It requires students to make a decision when only temporal constraints have to be satisfied. Students can use the boundary conditions on times when Fred, Stanley, and Isaac can see movies, match these against the showing times for "Children in the Net" in the table and determine the correct answer. A correct performance on *Cinema Outing,* Question 2 corresponds to Level 1 on the PISA problem-solving proficiency scale, as students only need to understand and check some information that is easily retrievable from the problem statement.

This problem is about finding a suitable time and date to go to the cinema.

Isaac, a 15-year-old, wants to organise a cinema outing with two of his friends, who are of the same age, during the one-week school vacation. The vacation begins on Saturday, 24th March and ends on Sunday, 1st April.

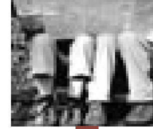
Isaac asks his friends for suitable dates and times for the outing. The following information is what he received.

Fred: "I have to stay home on Monday and Wednesday afternoons for music practice between 2:30 and 3:30."

Stanley: "I have to visit my grandmother on Sundays, so it can't be Sundays. I have seen Pokamin and don't want to see it again."

Isaac's parents insist that he only goes to movies suitable for his age and does not walk home. They will fetch the boys home at any time up to 10 p.m.

Isaac checks the movie times for the vacation week. This is the information that he finds.



TIVOLI CINEMA

Advance Booking Number: 01924 423000
24 hour phone number: 01924 420071
Bargain Day Tuesdays: All films $3
Films showing from Fri 23rd March for two weeks:

Children in the Net 113 mins 14:00 (Mon-Fri only) 21:35 (Sat/Sun only)	Suitable only for persons of 12 years and over	**Pokamin** 105 mins 13:40 (Daily) 16:35 (Daily)	Parental Guidance. General viewing, but some scenes may be unsuitable for young children
Monsters from the Deep 164 mins 19:55 (Fri/Sat only)	Suitable only for persons of 18 years and over	**Enigma** 144 mins 15:00 (Mon-Fri only) 18:00 (Sat/Sun only)	Suitable only for persons of 12 years and over
Carnivore 148 mins 18:30 (Daily)	Suitable only for persons of 18 years and over	**King of the Wild** 117 mins 14:35 (Mon-Fri only) 18:50 (Sat/Sun only)	Suitable for persons of all ages

CINEMA OUTING – Question 1

Taking into account the information Isaac found on the movies, and the information he got from his friends, which of the six movies should Isaac and the boys consider watching?

Circle "Yes" or "No" for each movie.

Movie	Should the three boys consider watching the movie?
Children in the Net	Yes / No
Monsters from the Deep	Yes / No
Carnivore	Yes / No
Pokamin	Yes / No
Enigma	Yes / No
King of the Wild	Yes / No

Response Coding guide for CINEMA OUTING Question 1

Full Credit

Code 2: Yes, No, No, No, Yes, Yes, in that order.

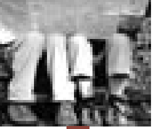
Partial Credit

Code 1: One incorrect answer.

No Credit

Code 0: Other responses.

Code 9: Missing.



CINEMA OUTING – Question 2

If the three boys decided on going to "Children in the Net", which of the following dates is suitable for them?

A. Monday, 26th March

B. Wednesday, 28th March

C. Friday, 30th March

D. Saturday, 31st March

E. Sunday, 1st April

Response Coding guide for CINEMA OUTING Question 2

Full Credit

Code 1: C. Friday, 30th March.

No Credit

Code 0: Other responses.

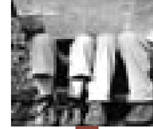
Code 9: Missing.

System analysis and design – the **CHILDREN'S CAMP** problem

Context: Community/Leisure
Levels: Level 2 (partial credit) and Level 3 (full credit)
PISA scale score: 529 (partial credit) and 650 (full credit)

Children's Camp is an example of a *system analysis and design* problem. Students have to understand the various constraints and their interrelationships, and design a solution that complies with them. This problem presents students with a statement about the context of a summer camp, lists of adult and child participants, and a set of boundary constraints that must be satisfied in the assignment of participants to the different dormitories at the camp. Full credit on this problem corresponds to proficiency Level 3. A correct solution requires students to combine different pieces of information about both the age and gender of the individuals involved. The students must arrange a match between the characteristics of the adults and children involved, and assign individuals to dormitories taking into account the capacities of the dormitories with respect to the number and gender of the children participating.

While a certain amount of trial and error can be used in working through the first phases to understand the problem, the successful solution requires students to monitor and adjust partial solutions relative to a number of interrelated conditions. A correct solution requires careful communication that details an appropriate number of the correctly matched students with an adult counsellor for each of the cabin dormitories. Students must work with several interrelated conditions and continually cross check until they have a solution that satisfies the constraints given. To do this, they must constantly shift between the desired state, the constraints, and the current status of their emerging solution. This requirement to manage the interactions simultaneously with the development of a unique solution is what makes the problem a Level 3 task.



The Zedish Community Service is organising a five-day Children's Camp. 46 children (26 girls and 20 boys) have signed up for the camp, and 8 adults (4 men and 4 women) have volunteered to attend and organise the camp.

Table 1. **Adults**

Mrs Madison
Mrs Carroll
Ms Grace
Ms Kelly
Mr Stevens
Mr Neill
Mr Williams
Mr Peters

Table 2. **Dormitories**

Name	Number of beds
Red	12
Blue	8
Green	8
Purple	8
Orange	8
Yellow	6
White	6

Dormitory rules:

1. Boys and girls must sleep in separate dormitories.
2. At least one adult must sleep in each dormitory.
3. The adult(s) in a dormitory must be of the same gender as the children.

CHILDREN'S CAMP – Question 1

Dormitory Allocation

Fill the table to allocate the 46 children and 8 adults to dormitories, keeping to all the rules.

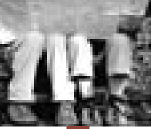
Name	Number of boys	Number of girls	Name(s) of adult(s)
Red			
Blue			
Green			
Purple			
Orange			
Yellow			
White			

Response Coding guide for CHILDREN'S CAMP Question 1

Full Credit

Code 2: 6 conditions to be satisfied

- Total girls = 26
- Total boys = 20
- Total adults = four female and four male



- Total (children and adults) per dormitory is within the limit for each dormitory
- People in each dormitory are of the same gender
- At least one adult must sleep in each dormitory to which children have been allocated

Partial Credit

Code 1: One or two conditions (mentioned in Code 2) violated. Violating the same condition more than once will be counted as ONE violation only.

- Forgetting to count the adults in the tally of the number of people in each dormitory
- The number of girls and the number of boys are interchanged (number of girls = 20, number of boys = 26), but everything else is correct. (Note that this counts as two conditions violated.)
- The correct number of adults in each dormitory is given, but not their names or gender. (Note that this violates both condition 3 and condition 5.)

No Credit

Code 0: Other responses.

Code 9: Missing.

Trouble shooting – the **IRRIGATION** problem

Context: Society

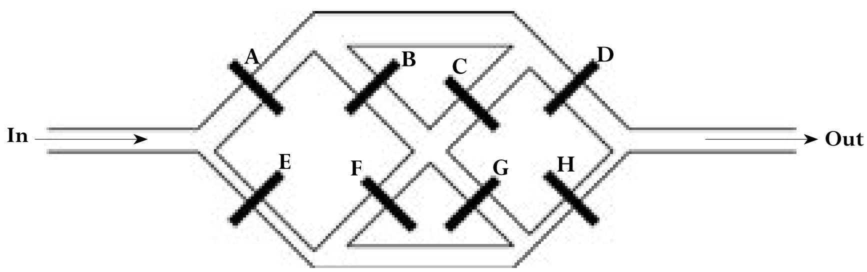
Levels: Level 1 (*Irrigation,* Question 1), Level 2 (*Irrigation,* Question 2 and *Irrigation,* Question 3)

PISA scale score: 497 (*Irrigation,* Question 1), 544 (*Irrigation,* Question 2) and 532 (*Irrigation,* Question 3)

Irrigation is an example of a *trouble-shooting* item. This problem presents students with a system of gates and canals, which provides means of distributing water across a network described by a pictorial diagram. *Irrigation,* Question 1 measures whether students understand the problem and how the gates in the irrigation network operate. Students proficient at Level 1 will typically answer correctly, as the task only requires the students to set the gates and then check if there is a path by which water can flow through the system. Students merely need to make a one-to-one transformation of the data from the table to the diagram and then trace it to see if there is a path from the inflow point to the outlet.

Below is a diagram of a system of irrigation channels for watering sections of crops. The gates A to H can be opened and closed to let the water go where it is needed. When a gate is closed no water can pass through it.

This is a problem about finding a gate which is stuck closed, preventing water from flowing through the system of channels.



Michael notices that the water is not always going where it is supposed to.

He thinks that one of the gates is stuck closed, so that when it is switched to open, it does not open.

IRRIGATION – Question 1

Michael uses the settings given in Table 1 to test the gates.

Table 1. **Gate Settings**

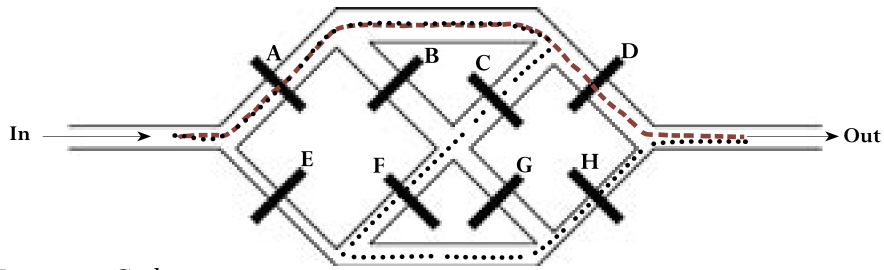
A	B	C	D	E	F	G	H
Open	Closed	Open	Open	Closed	Open	Closed	Open

With the gate settings as given in Table 1, **on the diagram below** draw all the possible paths for the flow of water. Assume that all gates are working according to the settings.

Response Coding guide for IRRIGATION Question 1

Full Credit

Code 1: Flow paths as shown below:



Response Coding notes:

Ignore any indications of the directions of flow.

Note that the response could be shown in the diagram provided, or in Figure A, or in words, or with arrows.

No Credit

Code 0: Other responses.

Code 9: Missing.

IRRIGATION – Question 2

Michael finds that, when the gates have the Table 1 settings, no water flows through, indicating that at least one of the gates set to open is stuck closed.

Decide for each problem case below whether the water will flow through all the way. Circle "Yes" or "No" in each case.

Problem Case	Will water flow through all the way?
Gate **A** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No
Gate **D** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No
Gate **F** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No

Response Coding guide for IRRIGATION Question 2

Full Credit

Code 1: No, Yes, Yes, in that order.

No Credit

Code 0: Other responses.

Code 9: Missing.

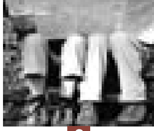
IRRIGATION – Question 3

Michael wants to be able to test whether **gate D** is stuck closed.

In the following table, show settings for the gates to test whether **gate D** is stuck closed when it is set to open.

Settings for gates (each one open or closed)

A	B	C	D	E	F	G	H



Response Coding guide for IRRIGATION Question 3

Full Credit

Code 1: A and E are not both closed. D must be open. H can only be open if water cannot get to it (*e.g.* other gates are closed preventing water from reaching H). Otherwise H must be closed.

- H closed, all other gates open

No Credit

Code 0: Other responses.

Code 9: Missing.

The second problem, *Irrigation,* Question 2, requires student performances typically associated with Level 2 problem solvers. Such students have to understand and trouble shoot the mechanism, in this case the system of gates and canals when the gates are set as given in the first problem, to locate the potential problem when water does not flow through the system. This requires the students to keep in mind the representation and then apply deductive and combinatorial reasoning in order to find a solution.

Similarly, *Irrigation,* Question 3 is a Level 2 problem because it requires students to handle several interconnected relationships at once, moving between the gate settings and possible flow patterns to ascertain whether a particular gate setting will result in water flowing or not flowing through Gate D.

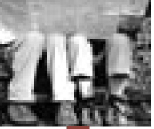
The three items above illustrate the three problem types at various levels of difficulty.

To summarise, these three items provide one example of each of the three problem types. In the *decision-making* problem students need to understand the given information, identify the relevant alternatives and the constraints involved, construct or apply external representations, select the best solution from a set of given alternatives and communicate the decision. In the *system analysis and design* problem students need to understand the complex relationships among a number of interdependent variables, identify their crucial features, create or apply a given representation, and design a system so that certain goals are achieved. Students also need to check and evaluate their work through the various steps along the way to an analysis or design. In the *trouble-shooting* problem students need to diagnose the problem, propose a solution and execute this solution. Students must understand how a device or procedure works, identify the relevant features for the task at hand and create a representation.

The percentage of students at each proficiency level of problem solving

Students at each proficiency level have at least a 50 per cent chance of solving problems at that level.

Figure 2.2 also shows where each item from the three problem units presented above is located on the PISA problem-solving scale. A student who scores 468 on this scale is likely to be able to answer *Cinema Outing,* Question 2 correctly. To be precise, students have a 62 per cent chance of answering correctly a task ranked at their point score. This is the criterion used throughout PISA, and has



been set in order to meet another condition: Each student is assigned to the highest level for which they would be expected to answer correctly the majority of assessment items. Thus, for example, in a test composed of items spread uniformly across Level 2 (with difficulty ratings of 499 to 592 score points), all students assigned to that level would expect to get at least 50 per cent of the items correct. Someone at the bottom of the level (scoring 499 points) would be expected to get close to 50 per cent of the items correct; someone in the middle or near the top of the level would get a higher percentage of items correct. For this to be true, a student scoring 499 points needs to have a 50 per cent chance of completing an item in the middle of level 3 and thus have a greater than 50 per cent chance of getting right an item rated at their score, 499 points. This latter probability needs to be 62 per cent to fulfil these conditions

Country performance can be summarised in terms of how many students are proficient at least at Level 3, Level 2 and Level 1.

Figure 2.3 and Table 2.1 classify students in participating countries by their highest level of problem-solving proficiency (note that a student proficient at Level 2, for example, is also proficient at Level 1). The percentage of students at or below Level 1 appears below the horizontal axis and the percentage at or above Level 2 appears above the same line. This shows at a glance how many students have higher level problem-solving skills compared to only basic problem-solving skills in each country. Note that this divide also corresponds approximately to how many students are above or below the OECD average in terms of problem-solving performance.

In some countries most students can solve relatively complex problems, while in others few can…

It is clear that in these terms country results vary greatly, from some countries where the great majority of students can solve problems at least at Level 2, to others where hardly any can. At the same time, the variation within countries in problem-solving ability is much larger. For example, in the majority of OECD countries, the top 10 per cent of students are proficient at Level 3, but the bottom 10 per cent of students are not proficient at Level 1 (Table 2.1).

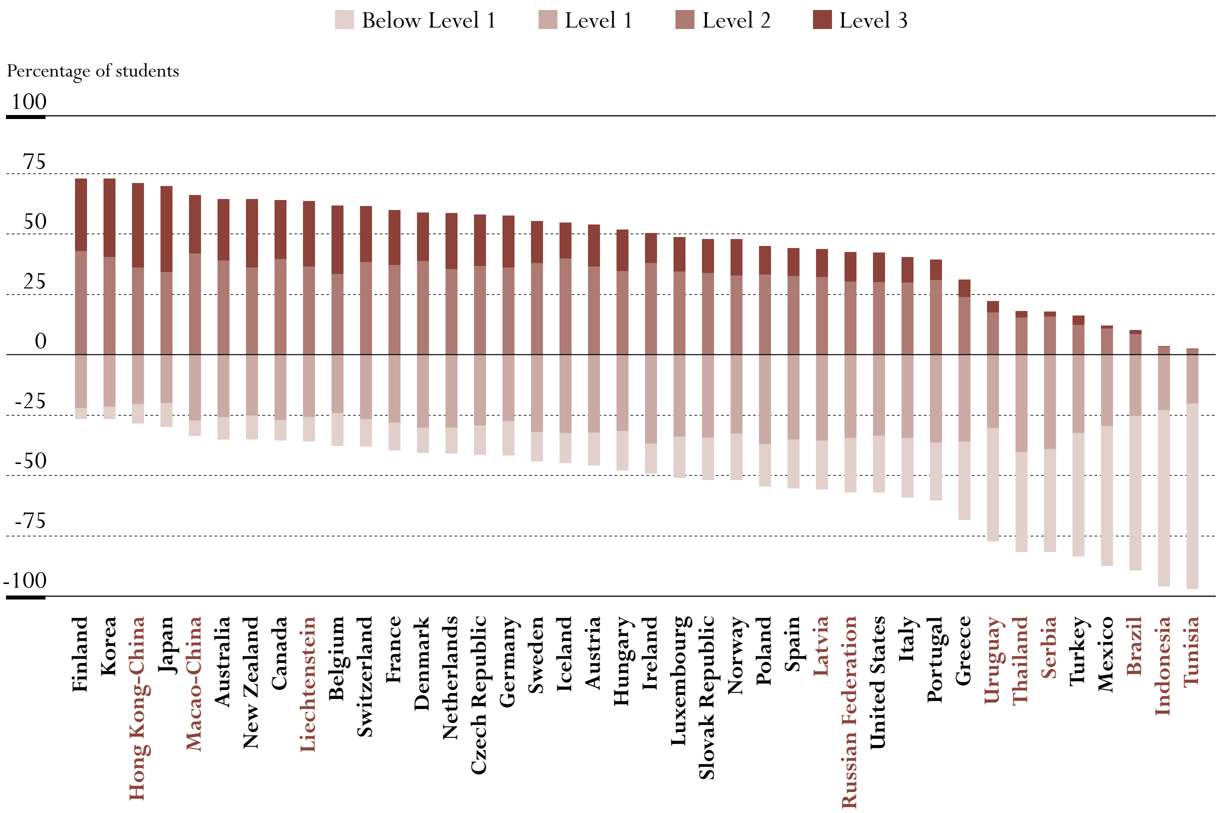
…with the proportion varying from above seven in ten students to below one in 20.

On average, about half of the students in OECD countries score at Level 2 or above. The national percentages of students at Level 2 or above range from 70 per cent or more in Finland, Japan, Korea, and the partner country Hong Kong-China, to less than 5 per cent in the partner countries Indonesia and Tunisia. Figure 2.3 also shows that more than a third of the students in Japan and the partner country Hong Kong-China perform at Level 3. In 26 OECD countries and five partner countries between 30 and 43 per cent of students are proficient at Level 2, but in eight PISA countries below 20 per cent of students are proficient at this level.

In most countries, more than one student in ten are unable to solve basic problems at Level 1, and in five countries over half the students are unable to do so.

The percentage of students with a low proficiency profile (unable to solve Level 1 problems) ranges from over half of all participating students in Mexico and Turkey, as well as in the partner countries Brazil, Indonesia and Tunisia, to below 10 per cent in Australia, Canada, Finland, Korea and the partner countries Hong Kong-China and Macao-China. There are comparatively high proportions of students with weak problem-solving skills in other OECD countries also: In Italy, Portugal and the United States, nearly a quarter fall

Figure 2.3 ■ **Percentage of students at each level of proficiency on the problem-solving scale**



Countries are ranked in descending order of percentage of 15-year-olds in Levels 2 and 3.
Source: OECD PISA 2003 database, Table 2.1.

below Level 1, and in Greece nearly a third do. The percentage of students proficient at Level 1 varies from 21 per cent in Japan and the partner countries Hong Kong-China and Tunisia to 40 per cent in the partner country Thailand. Note, however, that in Japan the relatively small number of students at Level 1 is associated with the fact that nearly three-quarters of students are above Level 1, whereas in Tunisia over three-quarters are below Level 1.

Mean performance of countries

Along with the analysis of how students within countries are distributed across the various levels of proficiency in problem solving, there is interest in an overall measure of proficiency in problem solving. This can be achieved by estimating a mean problem-solving score for the country. This is shown in Figure 2.4.

An overall mean score can be calculated for each country, though this hides variations.

As discussed in Box 2.1, when interpreting mean performance, only those differences between countries that are statistically significant should be taken into account. The figure shows those pairs of countries where the difference in their mean scores is sufficient to say with confidence that the higher performance by sampled students in one country holds for the entire population of enrolled

Figure 2.4 shows where significant differences in these means can be detected.

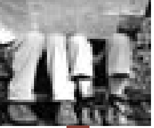
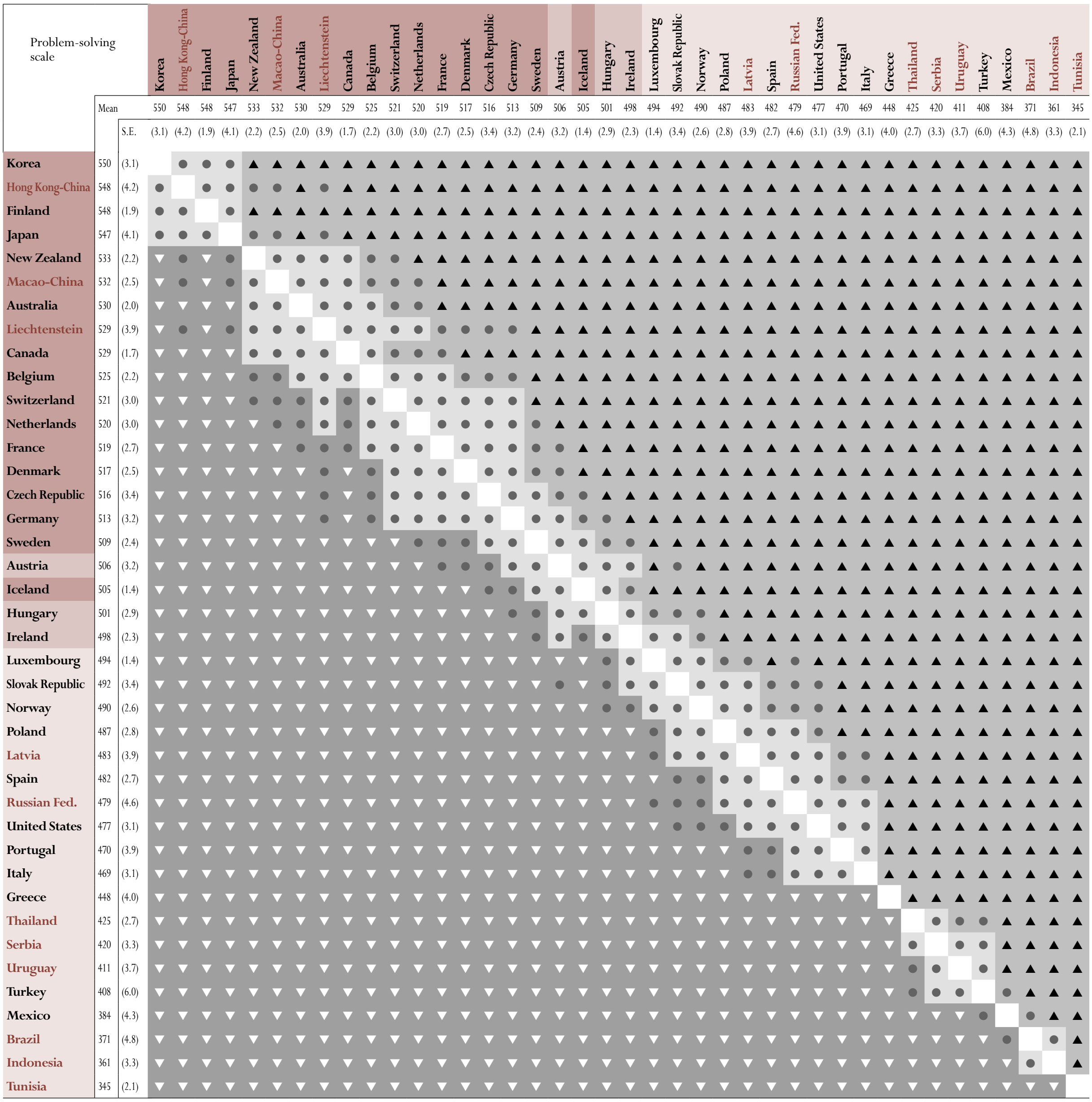


Figure 2.4 ■ Multiple comparisons of mean performance on the problem-solving scale



Range of rank*

		Korea	Hong Kong-China	Finland	Japan	New Zealand	Macao-China	Australia	Liechtenstein	Canada	Belgium	Switzerland	Netherlands	France	Denmark	Czech Republic	Germany	Sweden	Austria	Iceland	Hungary	Ireland	Luxembourg	Slovak Republic	Norway	Poland	Latvia	Spain	Russian Fed.	United States	Portugal	Italy	Greece	Thailand	Serbia	Uruguay	Turkey	Mexico	Brazil	Indonesia	Tunisia
OECD countries	Upper rank	1		1	1	4		4		4	6	7	7	7	8	8	10	12	13	14	15	17	18	18	19	20		22		23	24	24	27				28	29			
	Lower rank	3		3	3	6		7		7	9	12	12	13	13	14	15	16	17	17	19	19	21	22	22	23		24		25	26	26	27				28	29			
All countries	Upper rank	1	1	1	1	5	5	5	5	6	8	9	10	10	11	11	13	16	16	17	18	20	21	21	22	23	24	25	25	26	28	29	32	33	33	34	34	37	38	38	40
	Lower rank	4	4	4	4	8	9	10	11	10	12	15	15	16	16	17	18	19	20	20	22	22	24	26	26	27	29	29	30	30	31	31	32	34	35	36	36	37	39	39	40

* Because data are based on samples, it is not possible to report exact rank order positions for countries. However, it is possible to report the range of rank order positions within which the country mean lies with 95 per cent likelihood.

Instructions:

Read across the row for a country to compare performance with the countries listed along the top of the chart. The symbols indicate whether the average performance of the country in the row is lower than that of the comparison country, higher than that of the comparison country, or if there is no statistically significant difference between the average achievement of the two countries.

Without the Bonferroni adjustment:
- Mean performance statistically significantly higher than in comparison country
- No statistically significant difference from comparison country
- Mean performance statistically significantly lower than in comparison country

With the Bonferroni adjustment:
- ▲ Mean performance statistically significantly higher than in comparison country
- ● No statistically significant difference from comparison country
- ▽ Mean performance statistically significantly lower than in comparison country

- Statistically significantly above the OECD average
- Not statistically significantly different from the OECD average
- Statistically significantly below the OECD average

Source: OECD, PISA 2003 database.

Box 2.1 ■ **Interpreting sample statistics**

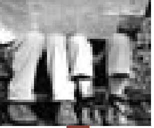
Standard errors and confidence intervals. The statistics in this report represent *estimates* of national performance based on samples of students rather than the values that could be calculated if every student in every country had answered every question. Consequently, it is important to know the degree of uncertainty inherent in the estimates. In PISA 2003, each estimate has an associated degree of uncertainty, which is expressed through a standard error. The use of confidence intervals provides a means of making inferences about the population means and proportions in a manner that reflects the uncertainty associated with sample estimates. Under the usually reasonable assumption of a normal distribution, and unless otherwise noted in this report, there is a 95 per cent chance that the true value lies within the confidence interval.

Judging whether populations differ. This report tests the statistical significance of differences between the national samples in percentages and in average performance scores in order to judge whether there are differences between the populations that the samples represent. Each separate test follows the convention that, if in fact there is no real difference between two populations, there is no more than a 5 per cent probability that an observed difference between the two samples will erroneously suggest that the populations are different as the result of sampling and measurement error. In the figures and tables showing multiple comparisons of countries' mean scores, multiple comparison significance tests are also employed that limit to 5 per cent the probability that the mean of a given country will erroneously be declared to be different from that of any other country, in cases where there is in fact no difference.

15-year-olds. A country's performance relative to that of the countries listed along the top of the figure can be seen by reading across each row. The colours indicate whether the average performance of the country in the row is either lower than that of the comparison country, not statistically significantly different, or higher. When making multiple comparisons, *e.g.* when comparing the performance of one country with that of all other countries, a more cautious approach is required: Only those comparisons indicated by the upward and downward pointing symbols should be considered statistically significant for the purpose of multiple comparisons.[1] Figure 2.4 also shows which countries perform above, at, or below the OECD average. Results from the United Kingdom were excluded from this and similar comparisons, because the data for England did not comply with the response rate standards which OECD countries had established to ensure that PISA yields reliable and internationally comparable data.

The top performers in problem solving are Finland, Japan and Korea and the partner country Hong Kong-China.

The top performing countries are Finland, Japan and Korea, as well as the partner country Hong Kong-China. These four countries perform indistinguishably well and are almost 50 score points, or around one-half of a proficiency level, ahead of the mean performance level for OECD countries, which is 500 score points. Other countries performing above this average are Australia, Belgium, Canada, the Czech Republic, Denmark, France, Germany, Iceland, the Netherlands,



New Zealand, Sweden and Switzerland and the partner countries Liechtenstein and Macao-China. Another three countries, Austria, Hungary and Ireland, are clustered around the OECD mean, but their performances are not statistically significantly different from it.[2] The remaining 19 countries all perform below the OECD average. These countries were Greece, Italy, Luxembourg, Mexico, Norway, Poland, Portugal, the Slovak Republic, Spain, Turkey and the United States, as well as the partner countries Brazil, Indonesia, Latvia, the Russian Federation, Serbia,[3] Thailand, Tunisia and Uruguay.

The distribution of problem-solving capabilities within countries

Within-country variation around the mean can be used to describe patterns of results…

While comparisons based on country means are useful in establishing the ordered list shown in Figure 2.4, they do little to describe the within-country variation in performance.

Figure 2.5 sheds further light on the performance distribution within countries. This analysis needs to be distinguished from the examination of the distribution of student performance across the PISA proficiency levels discussed above. Whereas the distribution of students across proficiency levels indicates the proportion of students in each country that can demonstrate a specified level of knowledge and skills, and thus compares countries on the basis of absolute benchmarks of student performance, the analysis below focuses on the relative distribution of scores, *i.e.* the gap that exists between students with the highest and the lowest levels of performance within each country. This is an important indicator of the equality of educational outcomes in mathematics.

…and a key indicator is the degree of performance variation within each country's distribution shown by the length of the bars in Figure 2.5.

The gradation bars in the figure show the range of performance in each country between the 5th percentile (the point below which the lowest-performing 5 per cent of the students in a country score) and the 95th percentile (the point below which 95 per cent of students perform or, alternatively, above which the 5 per cent highest-performing students in a country score). The density of the bar represents the proportion of students performing at the corresponding scale points. The solid, horizontal black line near the middle shows the mean score for each country (*i.e.* the subject of the discussion in the preceding section) and is located inside a shaded box that shows its confidence interval. The figure also indicates the 10th, 25th, 75th and 90th percentiles, *i.e.* the points above which 90, 75, 25 and 10 per cent of students perform. The data related to Figure 2.5 are in Table 2.2 in Annex B.

The figure shows that there is wide variation in overall student performance on the mathematics scale within countries. The middle 90 per cent of the population shown by the length of the bars exceeds by far the range between the mean scores of the highest and lowest performing countries. The wide range covered by the performance distribution in each country suggests that educational programmes, schools or teachers need to cope with a wide range of student knowledge and skills.

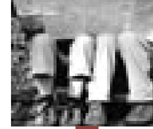
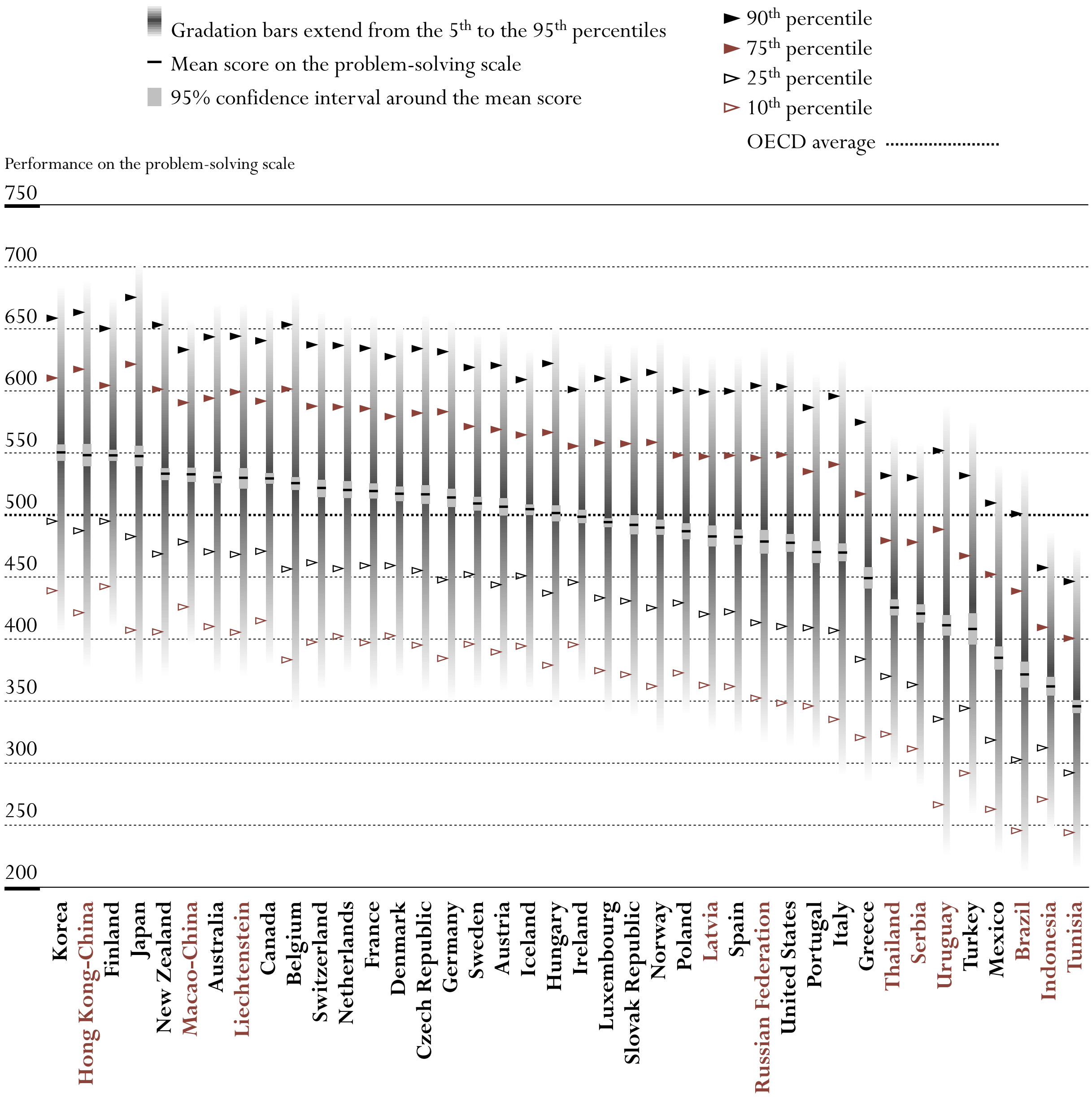


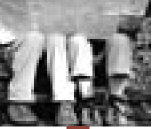
Figure 2.5 ■ **Distribution of student performance on the problem-solving scale**



Source: OECD PISA 2003 database, Table 2.2.

Compare, for example, the vertical bars for Belgium, a country with an above-average performance overall with those for Korea, the country with the highest mean score. The top of the bar and the first two identified lines (the 95th and 90th percentiles) are at similar points in these two countries, indicating that towards the top of the distribution students in Belgium perform at similar levels to their equivalents in Korea. However, when one looks further down the distribution, the performance of students in Belgium falls further below their peers than is the case for students in Korea. By the fifth percentile, students in Belgium are 64 score points, equivalent to two-thirds of a proficiency level,

For example, Belgium's highest performers do very well, as do those of Korea, but in the former there is much greater variation among weaker students.



behind students in Korea. This greater range of performance in Belgium is shown by a much longer bar, and the relatively wide range of performance in the bottom part of the student distribution is shown by the fact that the lower segments are wider than those at the top.

Other countries have distinctive patterns, with Finland and Korea, as well as the partner country Macao-China, doing best in having strong performance overall with low variation.

Compared to the other three top performing countries, Japan has more students performing at both the highest and lowest levels. The students in Finland and Korea, as well as in the partner country Macao-China, in contrast, have lower variation between the 5^{th} to 95^{th} percentiles points than other high-performing countries.

Patterns of within-country variation differ considerably across countries.

Overall, there is a great deal of variability in patterns of problem-solving capabilities of 15-year-olds in different countries. It is true that performance differences of students within each country are much greater than performance differences between countries: The difference between the means of the highest and lowest performing country (206 score points) is less than the range of performance between the 95^{th} and 5^{th} percentile points within each participating country. However, significant differences between countries in the average performance of students should not be overlooked and may raise questions about the future competitiveness of certain countries.

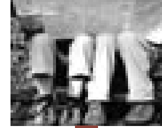
Implications for policy

The PISA cross-disciplinary assessment of problem solving was designed to assess the degree to which students in the OECD and the participating partner countries could solve problems situated in contexts that were not confined to one discipline and drew on students' knowledge from a variety of sources.

The design of the assessment placed particular emphasis on testing each student's ability to understand a problem situation, identify relevant information or constraints, represent possible alternatives or solution paths, select a solution strategy, solve the problem, check or reflect on the solution, and communicate the solution and reasoning behind it.

One in six OECD students, but much more in some countries, lack basic problem-solving skills.

The analysis of student performance on this assessment shows that about one in six students in the OECD area are only able to work in highly structured and straightforward settings, where they can deal with information available from direct observation or from very simple inferences. They are generally unable to analyse situations or solve problems that call for anything other than the direct collection of information, and are therefore characterised as *weak or emergent problem solvers*. In four countries – Mexico and the partner countries Brazil, Indonesia and Tunisia – mean student performance falls at this level (mean performance is less than 405 score points). The four highest performing countries – Finland, Japan, Korea and the partner country Hong Kong-China – have between 5 and 10 per cent of their students performing below Level 1. In the OECD countries on average, 17 per cent of students perform below Level 1.



On average in the OECD countries 30 per cent of students are *basic problem solvers*. The mean performance of students in sixteen of the participating countries is at Level 1: Greece, Ireland, Italy, Luxembourg, Norway, Poland, Portugal, the Slovak Republic, Spain, Turkey and the United States and the partner countries Latvia, the Russian Federation, Serbia, Thailand and Uruguay. These students are consistently able to understand the nature of a problem and the relevant data associated with a problem's major features. In many situations, these 15-year-olds are able to make minor translations between ways in which the problem data might be represented. In addition, these students are generally capable of using the information to check a limited number of fairly direct statements related to the problem. However, the *basic problem solvers* are generally incapable of dealing with multi-faceted problems involving multiple data sources or requiring analytical reasoning with the information provided.

Basic problem solvers at Level 1 can deal with simple problems only…

On average in the OECD countries, 34 per cent of students are *reasoning, decision-making problem solvers*. Mean student performance in 20 countries participating in PISA 2003 is at Level 2: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Hungary, Iceland, Japan, Korea, the Netherlands, New Zealand, Sweden, Switzerland and the partner countries Hong Kong-China, Liechtenstein and Macao-China. The line between students performing at Level 1 and those performing at Level 2 is an important line of demarcation in terms of student competencies in problem solving. Students at Level 1 are limited, in general, to handling relatively straightforward problems that require the collection of information from readily available sources and making simple transformations of data from graphical forms or tabular forms to numerical forms for interpretation. *Basic problem solvers* are generally not capable of drawing data from multiple sources, comparing and contrasting these data and integrating the data into the development of a solution to a multifaceted problem. These are the very skills that are necessary in emergent workforce demands. New employee qualifications are focusing on the ability to deal with complexity, on communication skills, and on increased problem-solving capabilities (Green *et al.,* 1997; Lerman and Skidmore, 1999; Johnson, 2000; Steedman, 1999; Workbase, 2000).

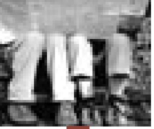
…and the ability to perform more complex tasks represents an important performance threshold between Level 1 and Level 2…

As 15-year-olds develop the problem-solving skills associated with Levels 2 and 3 of the PISA problem-solving scale, they have increased opportunities for employment and the ability to successfully participate in a rapidly changing world. They also exhibit the problem-solving knowledge and skills associated with enfranchised citizenship. These skills are marked by the problem-solving actions and outcomes described in the following two levels.

…with potential real-world consequences.

Students performing at Level 2 exhibit the capacity to apply analytical reasoning skills to solve problems involving decision making that requires comparisons of multiple alternatives. In doing so, these *reasoning, decision-making problem solvers* are able to handle a variety of representations of related information and use them to select the best of several alternatives in a variety of contexts. 15-year-olds at Level 2 are also capable of drawing inferences in settings involving applications of deductive, inductive, and combinatorial reasoning.

Level 2 is associated with crucial thinking skills.



Fewer students can perform the most complex tasks at Level 3 than at the previous levels, but in some countries over a quarter can do so.

The highest identified level of problem solving, Level 3, includes student work that reflects not only the ability to confront and derive a solution to a problem, but also the capability to reflect on and use information about underlying relationships found in problem situations. Students at this level can construct their own representations of problems from pieces of information and then in systematic ways solve the problems and communicate their findings to others. *Reflective, communicative problem solvers* are capable of handling a greater number of variables, of handling time and sequential relationships, and a variety of other problem-specific constraints. None of the participating countries have their student mean performance in problem solving at Level 3. However, in Australia, Belgium, Canada, Finland, Japan, Korea and New Zealand and the partner countries Hong Kong-China and Liechtenstein 25 per cent or more of the students are capable of solving problems at this highest level. The four countries with the highest overall averages in problem solving – Finland, Japan, Korea, and the partner country Hong Kong-China – all have 30 per cent or more of their students at Level 3. On average in the OECD countries 18 per cent of students are *reflective, communicative problem solvers*.

However, on average in the OECD countries, half of students lack the important ability to solve problems that are more complex than the basic problems at Level 1.

If one considers the percentage of students performing at either Level 2 or Level 3 on the problem-solving scale to be one indicator of how well prepared 15-year-olds are for productive participation in an emerging 21st century society, the majority of students are prepared in only 22 of the 40 participating countries with comparable data. These data, combined with data from students' performances in other content areas, indicate that student literacy levels fail to meet expectations in many of the countries participating in PISA 2003. Such data also add to the emerging data on the spread of wealth and opportunity in countries worldwide.

Notes

1. Although the probability that a particular difference will falsely be declared to be statistically significant is low (5 per cent) in each single comparison, the probability of making such an error increases when several comparisons are made simultaneously. It is possible to make an adjustment for this which reduces to 5 per cent the maximum probability that differences will be falsely declared as statistically significant at least once among all the comparisons that are made. Such an adjustment, based on the Bonferroni method, has been incorporated into the multiple comparison charts in this volume, as indicated by the arrow symbols.
2. Note that while Iceland has a lower mean performance level than Austria, its standard error is also less than that of Austria. This leads to Iceland being statistically significantly above the OECD average, while Austria is found to be not statistically significantly different from the OECD average.
3. For the country Serbia and Montenegro, data for Montenegro are not available. The latter accounts for 7.9 per cent of the national population. The name "Serbia" is used as a shorthand for the Serbian part of Serbia and Montenegro.

Student Performance in Problem Solving Compared with Performance in Mathematics, Reading and Science

Introduction

This chapter looks at how problem solving was assessed and what the results were, in relation to the other parts of PISA.

This chapter looks further at the results of the PISA assessment of problem solving, and compares them with results from the mathematics, science and reading assessments.

Problem-solving framework and test development

Problem solving differs from mathematics, reading, and science in that it is not a traditional school subject. However, problem-solving skills are required in curricular subjects as well as in non-curricular areas. There are two main features that distinguish problem solving from the PISA assessments of mathematics, science and reading – the first relates to an emphasis on problem-solving processes, the second relates to the much lower levels of content knowledge required for the tasks.

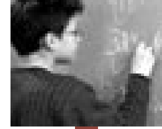
Emphasis on problem-solving processes

Each section of the assessment required students to show that they could solve as well as recognise a problem…

Chapter 2 described the processes that underlie PISA problem-solving tasks. These relate to understanding the problem, characterising the problem, representing the problem, solving the problem, reflecting on the problem and communicating the problem solution. While not every problem-solving task in PISA involved all of the processes listed, it was established that each problem-solving unit should test students on most of these processes. A problem-solving unit refers to the description of a problem-solving situation followed by a number of individual questions about it, which are referred to as items. While some items within a unit may test only the understanding of the stated problem, it was deemed important that each problem-solving unit, at some point, require students to demonstrate their ability in actually solving a problem. For example, if a unit is of the decision-making problem type, then students are asked to make some decisions. If a unit is of the system analysis and design type, then students are asked to actually analyse the problem or design a solution. From this point of view, the problem-solving units differ from tasks in reading in that reading comprehension, while an important part of problem solving, is not the only process tested within a unit. Similarly, if a task requires some mathematical operations, these mathematical skills are not the only skills tested in a problem-solving unit.

…thus assessing problem solving as a whole process, not just the component skills in isolation.

Consequently, problem-solving units differ from those in the other assessment areas in that the units test for a majority of the problem-solving processes, not only for isolated processes such as reading comprehension, mathematical manipulations, or communication. The problem-solving units included in Chapter 4 show that each unit contains at least one item that leads to a decision or to an analysis or design of solutions. However, individual items that test the extent to which students understand a problem situation are also informative in showing whether the failure to arrive at a solution is due to a lack of understanding of the problem given. Thus, there is a hierarchy of problem-solving processes, in that understanding the problem is at the most basic level of the processes,



without which no solutions can be reached. But to be proficient in problem solving, other skills relating to representation and analysis of problems are essential.

Low content requirements

To avoid this being a reading test, the amount and difficulty of reading required was limited…

Given that PISA 2003 was a paper-and-pencil test, it was inevitable that written text was required to state the problem situation. Consequently, as noted above, reading comprehension is a pre-requisite for problem solving. However, written texts were kept to a minimum in the problem-solving units. Of the ten problem-solving units in the test (presented in Chapter 4), only one unit is solely text-based in both the stimulus and in the items (*Freezer*). The other units use texts as well as diagrams and tables to convey the information, and to elicit answers. Two units (*Design by Numbers©* and *Transit System*) have very little text involved, as both the stimulus and answer format involve diagrams or figures. The other units involve both texts and tables or diagrams, with minimal amount of reading load. Overall, the test designers were conscious of the amount of reading involved, and kept the level of text difficulty as low as possible, with nothing beyond everyday use of language.

…and the problems required only simple mathematical skills and no scientific knowledge.

Similarly, where mathematical manipulations were required, the tasks were limited to very simple mathematical operations (*Energy Needs* and *Holiday*). In contrast, in the PISA mathematics assessment, every item has mathematical content, most of which is beyond the level of simple, straightforward operations. In relation to science, no problem-solving item in the PISA 2003 assessment involved either scientific content or context.

Consequently, the items in the problem-solving assessment differ from those in the other three assessment areas in that there is little or no curriculum content required in order to correctly complete the problems as presented.

The key skills tested in problem solving

The key skill needed to solve problems is analytical reasoning…

Given that the assessment of content knowledge is minimised, what are the main skills tested in the problem-solving assessment? The main cognitive skills tested relate to analytical reasoning. In most items, students are required to organise and analyse information, and derive solutions satisfying given constraints. Many problem situations in everyday life do not involve high levels of curriculum content knowledge. Instead, they involve the ability to reason, and the ability to approach problems in systematic ways such as elimination and enumeration. These are the key skills tested in the problem-solving assessment. From this point of view, problem solving fills a gap by testing the use of these skills in problem situations not already covered in the other three assessment areas.

…which is also an important requirement for mathematics tasks.

Even though problem-solving items do not involve mathematical content knowledge, other than simple arithmetic in two items, it is still expected that there will be a high correlation between student performance in mathematics and problem solving. This is because mathematics also requires a high level of analytical reasoning skills, particularly for 15-year-olds, who are generally beyond the level of mastery of basic skills in mathematics (*e.g.* Carroll, 1996).

Figure 3.1 ■ Analysis of two dominant factors in student performance on the problem-solving, reading and mathematics items

A. Factor 1 has a higher loading than Factor 2

Problem-solving item

Mathematics item

Number of items loading higher on Factor 1 than on Factor 2	Item[1]	Factor loadings[2]	
		Factor 1	Factor 2
0.350 to 0.400	HOLIDAY – Question 2	0.393	
11 mathematics items	SKATEBOARD – Question 13	0.391	
1 problem-solving item	NUMBER CUBES – Question 3	0.368	
0.300 to 0.349	IRRIGATION – Question 3	0.346	
19 mathematics items	ROBBERIES – Question 15	0.345	
7 problem-solving items	SKATEBOARD – Question 12	0.335	
	WALKING – Question 5	0.334	
	COURSE DESIGN – Question 1	0.328	
	IRRIGATION – Question 2	0.321	
	LIBRARY SYSTEM – Question 2	0.318	
	IRRIGATION – Question 1	0.313	
	HOLIDAY – Question 1	0.310	
	ENERGY NEEDS – Question 2	0.303	
	WALKING – Question 4	0.301	
0.250 to 0.299	TEST SCORES – Question 16	0.298	
20 mathematics items	SKATEBOARD – Question 14	0.298	
3 problem-solving items	TRANSIT SYSTEM – Question 1	0.292	
	DESIGN BY NUMBERS© – Question 3	0.285	
	EXPORTS – Question 18	0.281	
	CARPENTER – Question 1	0.275	
	CHILDREN'S CAMP – Question 1	0.271	
0.200 to 0.249	CINEMA OUTING – Question 1	0.234	
16 mathematics items	DESIGN BY NUMBERS© – Question 1	0.234	
4 problem-solving items	LIBRARY SYSTEM – Question 1	0.232	
	STAIRCASE – Question 2	0.213	
	DESIGN BY NUMBERS© – Question 2	0.207	
	GROWING UP – Question 8	0.201	0.198
0.150 to 0.199	CINEMA OUTING – Question 2	0.189	
9 mathematics items	FREEZER – Question 1	0.188	
4 problem-solving items	EXPORTS – Question 17	0.176	
	FREEZER – Question 2	0.170	
	ENERGY NEEDS – Question 1	0.157	

B. Factor 2 has a higher loading than Factor 1

Problem-solving item

Mathematics item

Number of items loading higher on Factor 2 than on Factor 1	Item[1]	Factor loadings[2]	
		Factor 1	Factor 2
0.500 to 0. 650 9 reading items			
0.300 to 0.499 4 reading items			
0.200 to 0.299 11 reading items 3 mathematics items	EXCHANGE RATE – Question 10 GROWING UP – Question 7 EXCHANGE RATE – Question 9	0.201 0.181 0.165	0.227 0.223 0.217
0.100 to 0.199 1 reading item 6 mathematics items	GROWING UP – Question 6 EXCHANGE RATE – Question 11	0.182 0.193	0.196 0.193

1. Chapter 4 presents all problem-solving items. Mathematics items listed in this figure are presented in *Learning for Tomorrow's World – First Results from PISA 2003* (OECD, 2004a).
2. Factor loadings of less than 0.1 are not listed.
Source: OECD PISA database 2003, Table 3.1.

In order to better understand the cognitive demands of the problem-solving items, an exploratory analysis was carried out to identify patterns in student responses across PISA that suggest which groups of tasks are being influenced by certain common factors. This factor analysis was carried out using a random selection of 500 students from each OECD country participating in the PISA survey. Details of how it was conducted are given in Annex A2.

The overlap in such cognitive requirements can be analysed through patterns of student responses.

The results of the exploratory factor analysis suggest that different factors were influencing students' performance in reading and in mathematics, with problem-solving responses more closely associated with the mathematics factor. The full results of the factor analysis are provided in Table 3.1, Annex B. The main results for all problem-solving items as well as those mathematics items published in *Learning for Tomorrow's World – First Results from PISA 2003* are presented in Figure 3.1. The analysis identified two presumed factors. The factor loadings shown in these figures indicate the strength of association with the two presumed factors. Figure 3.1, part A, shows all items that loaded higher on the first presumed factor than on the second presumed factor. (Loadings less than 0.1 in absolute values were omitted from the table and figures.) No reading items loaded higher on the first presumed factor than on the second presumed factor, however in total 75 mathematics items and 19 problem-solving items did. Bearing in mind that problem-solving items do not contain mathematical content other than simple arithmetic, this suggests that analytical reasoning is the skill identified in the first factor

Mathematics and problem solving seem to be closely linked…

…by the common analytical requirements of most items.

Figure 3.1, part B, shows items that loaded higher on the second presumed factor than on the first presumed factor. This was the case for 25 reading items and nine mathematics items. However, no problem-solving items loaded higher on the second factor than on the first factor and that holds even for those items that only tested the understanding of the problem. The second factor is therefore more clearly linked to reading items. The items with high loadings on the first factor (identified as the analytical reasoning factor) include Questions 1 and 2 of the *Holiday* unit, Questions 1, 2 and 3 of the *Irrigation* unit, Question 1 of the *Course Design* unit, Question 2 of the *Library System* unit, and Question 2 of the *Energy Needs* unit. These items all require high levels of analysing, reasoning, organising, checking and evaluating skills.

Problem-solving items requiring less analysis load lower…

Problem-solving items that load relatively low on the first factor have characteristics that do not require a high level of analytical reasoning. For example, Questions 1 and 2 of the *Freezer* unit require an analysis of an everyday trouble-shooting situation, but unlike many other problem-solving questions do not involve the manipulation of parameters to satisfy constraints and specifications. Another example is Question 1 of the *Energy Needs* unit, which involves only looking up information in a table. Question 2 of the *Cinema Outing* unit also loaded relatively low on the first factor, presumably because it involved looking up information with a limited number of constraints.

…as do mathematics items requiring only basic computation.

Figure 3.1 not only reveals the nature of the problem-solving items, but also the nature of the mathematics items. The mathematics items loading high on the first factor require skills in enumeration, combinatorial reasoning and analytical reasoning. In contrast, Question 17 of the *Exports* unit had a relatively low loading on the first factor and required students to simply look up information in a graph. What the public would normally regard as mathematics – such as doing arithmetic and basic operations – was, in fact, not what the PISA mathematics assessment focused on. Rather, it tested the ability to analyse and reason using mathematics.

However, other factors not explored here could also be important.

It must be noted that the two-factor analysis presented in Table 3.1, Annex B, does not provide a full explanation of the cognitive demands of the tasks in the PISA assessment. The data shown in Annex A2 indicate that there are other factors that could provide important information about the items, for example, in relation to curriculum content.

Correlations between performance in reading, mathematics, science and problem solving

Individual students tend to obtain similar results across the three assessment areas, especially in mathematics and problem solving.

Having identified some of the differences between problem solving and the other three assessment areas, it is possible to analyse and interpret relationships between students' performances in the PISA assessment areas.

Figure 3.2 shows the latent correlations between the four PISA assessment areas. Latent correlations are direct estimates of the strength of the association between student abilities. The high figures for the latent correlations suggest that

students doing well in one area are likely to do well in other areas as well. The observed magnitudes of the latent correlations confirm the descriptions of the cognitive skills tested in the four assessment areas. Not surprisingly, problem solving correlates most highly with mathematics. The next highest correlation is with reading. The correlation between problem solving and science is somewhat lower which, once again, is not surprising given the lack of scientific content or context in the problem-solving items.

Figure 3.2 ■ **Latent correlations between the four assessment areas**

	Mathematics	Reading	Science
Mathematics			
Reading	0.77		
Science	0.83	0.83	
Problem solving	**0.89**	**0.82**	**0.80**

The correlation between problem solving and mathematics is also of about the same order of magnitude as the correlations among the four mathematics sub-scales.

Comparison between performances in mathematics and problem solving at the country level

Country scores in these two assessment areas can be compared…

For each country, the relative standing of the country's performance in mathematics and in problem solving can be compared. Given that the scores for mathematics and for problem solving were both scaled with a mean of 500 and a standard deviation of 100 among OECD countries, if a country has different mean scores for mathematics and for problem solving, this shows a difference in terms of how that country performs relative to the OECD average.

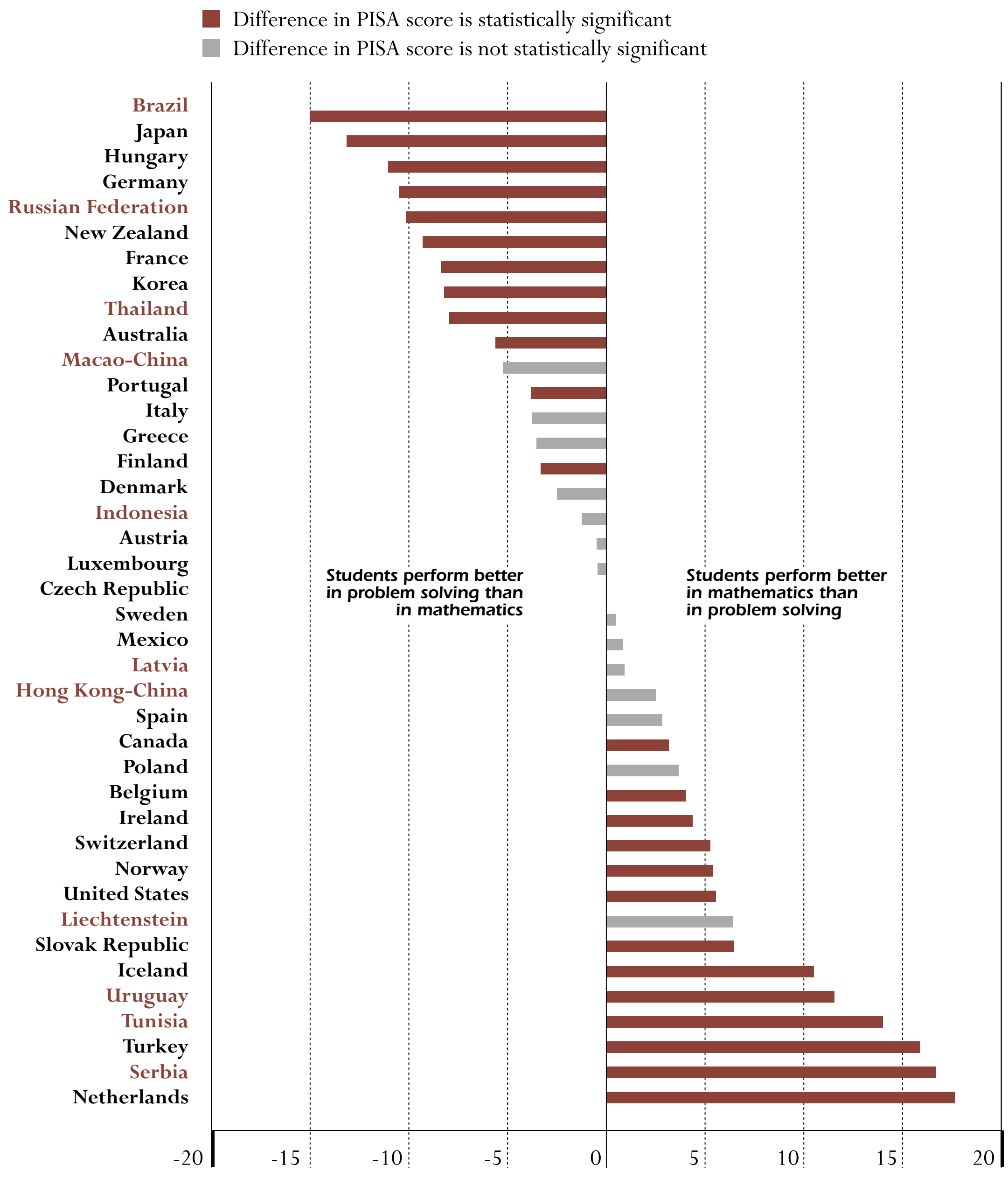
…giving an indication of whether student potential as shown by problem-solving skill is realised through the mathematics curriculum.

Taking a country's mean scores in mathematics and problem solving as the basis of comparison, if a country performs relatively better in mathematics than in problem solving, one can interpret this as showing that the students in the country have a better grasp of mathematics content as compared to other countries after accounting for the level of generic problem-solving skills of students. This may be an indication that mathematics instruction was particularly effective in the country. In contrast, if a country performs relatively better in problem solving, this may suggest that students have the potential to achieve better results in mathematics than that reflected in their current performance, since their level of generic problem-solving skills is relatively higher.

Although country scores are similar in the two assessment areas, some differences are noteworthy…

These differences are shown in Figure 3.3. In the Netherlands, Turkey and the partner country Serbia,[1] students scored on average between 16 and 18 points higher in mathematics than in problem solving, and there was a difference of at least 10 score points in Iceland and in the partner countries Tunisia and Uruguay. As explained above, this may indicate that mathematics instruction is relatively effective in helping students reach their potential. On the other hand,

Figure 3.3 ■ **Difference between student performance in mathematics and problem solving**

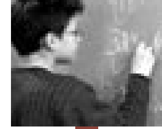


Source: OECD PISA 2003 database, Table 3.2.

...and in some cases, these show up in country rankings.

students scored between 10 and 15 points more on average in problem solving than in mathematics in Germany, Hungary and Japan, as well as in the partner countries Brazil and the Russian Federation. This may indicate that students have generic skills that may not be fully exploited by the mathematics curriculum.

The magnitude of these differences is at the most just under a quarter of a proficiency level on the problem-solving scale and about a third on the six-level mathematics scale. In some cases, however, there are significant differences in



country rankings on the two scales. For example, students in the Netherlands performed in the top five OECD countries in mathematics and are ranked between seventh and twelfth in problem solving. In mathematics, students in Hungary are ranked between nineteenth and twenty-third for OECD countries, but in problem solving they are ranked between fifteenth and nineteenth.

Implications for policy

Problem solving represents a competency area that is distinct from those based in a particular discipline.

The assessment of problem-solving skills in PISA 2003 takes the assessment further than in PISA 2000, in that problem solving is not a regular curriculum content area. Most work done with problem solving in the curriculum of schools is compartmentalised as problem solving in mathematics, in science, or in other subjects.

This distinction is clear despite an overlap in skills required, especially the reasoning skills needed for solving mathematical as well as general problems.

Even though the features of the tasks were quite different, there is still a strong relationship between students' performance levels in mathematics and reading and their performance in the problem-solving assessment. Analysis shows that two factors, a mathematics factor and a reading factor, account for a substantial amount of the variability in students' scores. However, the problem-solving items draw on aspects of reasoning that can be considered to be different to reasoning skills used in mathematics and reading. In fact, the reasoning skills identified as a factor in mathematics explain about 7.6 per cent of the variance in student performance in problem solving, while those identified as a factor in reading explain an additional 2.9 per cent of the variance.[2]

Stronger problem-solving competencies and weaker mathematics performance may indicate that the mathematics instruction provided does not fully exploit the potential of students.

The results in problem solving are clearly not identical to those in other assessment areas. For example, in some countries students do significantly better in either mathematics or problem solving, relative to the OECD average. In these cases, it is important to look more closely at the features of the curriculum and instructional styles that might contribute to these discrepancies. Such variation has its core perhaps in analytical reasoning abilities, as the problem-solving items in the PISA assessment made little use of calculation beyond very straightforward whole number skills. If a country performs relatively better in mathematics than in problem solving, this suggests that students in this country have a better grasp of mathematical content as compared to other countries after accounting for the level of generic problem-solving skills of students. This may be an indication that mathematics instruction was particularly effective in the country. In contrast, if a country performs relatively better in problem solving, this may suggest that students have the potential to achieve better results in mathematics than that reflected in their current performance, as their level of generic problem-solving skills is relatively higher.

Notes

1. For the country Serbia and Montenegro, data for Montenegro are not available. The latter accounts for 7.9 per cent of the national population. The name "Serbia" is used as a shorthand for the Serbian part of Serbia and Montenegro.
2. See Annex A2 for results from the factor analysis regarding the total variance explained.

4

Student Performance on the Problem-Solving Items

Introduction

This chapter describes all the questions used to assess students' problem solving capacity, showing how different skills were tested and how responses were marked.

The assessment of problem solving in PISA, like any assessment of student competencies, is centred on a framework and on the actual items used to assess what students can do. Chapter 1 discussed the framework; this chapter examines the problem-solving items. It looks at each of the 19 items used to assess problem solving, classified by their problem type: *decision making*, *system analysis and design*, and *trouble shooting*. The guide used by markers to code responses accompanies each item. Samples of student work are reproduced for many items with open-response formats. Where informative, these examples include work awarded both full credit and partial credit. This information is useful both to those interested in the international comparisons based on these items and to those involved in curriculum and teacher education activities in each of the countries participating in PISA 2003.

It covers the 19 items summarised in Figure 4.1…

The 19 items used in the assessment were clustered into ten units, which ranged from one item to three items in length. Figure 4.1 shows for each unit exactly how many items were used and the type of problem posed, as well as what kind of answer was required for each item.

…which used several forms of questions…

As in PISA 2000, the assessment instruments in PISA 2003 comprised "units of assessment" – a series of texts followed by a number of questions on various aspects of each text, aiming to make tasks as close as possible to those encountered in the real world. The questions varied in format. Some questions required students to construct their own responses, either by providing a brief answer from a wide range of possible answers (short-response items) or by constructing a longer response (open-constructed response items), allowing for the possibility of divergent, individual responses and opposing viewpoints. Other parts of the test were based on students constructing their own responses, but based on a very limited range of possible responses (closed-constructed response items), which are scored as either correct or incorrect. The remaining items were asked in multiple-choice format, in which students either made one choice from among four or five given alternatives (multiple-choice items) or a series of choices by circling a word or short phrase (for example "yes" or "no") for each point of credit (complex multiple-choice items).

…and sometimes allowed students partial credit for not fully correct answers.

The following description presents each unit of items along with the ways in which credit could be awarded for each item, and what PISA scale score and proficiency level those marks correspond to. On 11 items, only one level of credit was possible – for a correct answer; on seven items, students could get a partial or a full credit mark; and on one item, two partial credit scores or a full credit score were possible. Thus, a total of 28 different score point levels could be given for the 19 items. Partial credit was provided for partially correct or less sophisticated answers, and all of these items were marked by experts. To ensure consistency in the marking process, many of the more complex items were marked independently by up to four markers and a subsample of student responses from each country was marked independently in at least two of the

participating countries. Finally, to verify that the marking process was carried out in equivalent ways across countries, an inter-country reliability study was carried out on a sub-set of items. The results show that very consistent marking was achieved across countries (for details on the marking process see Annex A7 and the *PISA 2003 Technical Report* [OECD, forthcoming]).

Figure 4.1 ■ **Problem-solving units and their characteristics**

Unit name	Item format
Decision-making units	
Energy Needs Q1	Closed-constructed response
Energy Needs Q2	Open-constructed response
Cinema Outing Q1	Multiple choice
Cinema Outing Q2	Multiple choice
Holiday Q1	Closed-constructed response
Holiday Q2	Open-constructed response
Transit System Q1	Open-constructed response
System analysis and design units	
Library System Q1	Closed-constructed response
Library System Q2	Open-constructed response
Design by Numbers Q1	Multiple choice
Design by Numbers Q2	Multiple choice
Design by Numbers Q3	Open-constructed response
Course Design Q1	Open-constructed response
Children's Camp Q1	Open-constructed response
Trouble-shooting units	
Irrigation Q1	Open-constructed response
Irrigation Q2	Multiple choice
Irrigation Q3	Open-constructed response
Freezer Q1	Multiple choice
Freezer Q2	Multiple choice

Decision-making units

The first items to be examined are those associated with *decision making*. These items present students with a situation requiring a decision and ask them to choose among alternatives under a set of conditions constraining the situation. Students have to understand the situation provided, identify the constraints, possibly translate the way in which the information is presented, make a decision based on the alternatives under the constraints given, check and evaluate the decision, and then communicate the required answer. The factors creating difficulty in decision-making problems are the number of constraints the student has to deal with in working through the information provided and the amount of restructuring a student has to do in sorting through the information along the way to developing a solution.

There are four *decision-making* units in the item set. Three of the units have two questions each and one unit consists of a single question.

ENERGY NEEDS

This unit asks two questions of students. The first problem, shown below, is about selecting suitable food to meet the energy needs of a person in Zedland. Success indicated that a student was able to look up needed information for solving a problem. This item's demands were below those associated with Level 1.

Daily energy needs recommended for adults

		Men	Women
Age (years)	**Activity level**	**Energy needed *(kJ)***	**Energy needed *(kJ)***
From 18 to 29	Light Moderate Heavy	10660 11080 14420	8360 8780 9820
From 30 to 59	Light Moderate Heavy	10450 12120 14210	8570 8990 9790
60 and above	Light Moderate Heavy	8780 10240 11910	7500 7940 8780

Activity level according to occupation

Light:	*Moderate:*	*Heavy:*
Indoor sales person Office worker Housewife	Teacher Outdoor salesperson Nurse	Construction worker Labourer Sportsperson

ENERGY NEEDS – Question 1

Mr David Edison is a 45-year-old teacher. What is his recommended daily energy need in kJ?

Answer: .... kilojoules.

Jane Gibbs is a 19-year old high jumper. One evening, some of Jane's friends invite her out for dinner at a restaurant. Here is the menu:

	MENU	**Jane's estimate of energy per serving** *(kJ)*
Soups:	Tomato Soup	355
	Cream of Mushroom Soup	585
Main courses:	Mexican Chicken	960
	Caribbean Ginger Chicken	795
	Pork and Sage Kebabs	920
Salads:	Potato salad	750
	Spinach, Apricot and Hazelnut Salad	335
	Couscous Salad	480
Desserts:	Apple and Raspberry Crumble	1380
	Ginger Cheesecake	1005
	Carrot Cake	565
Milkshakes:	Chocolate	1590
	Vanilla	1470

The restaurant also has a special fixed price menu.

Fixed Price Menu
50 zeds
Tomato Soup
Caribbean Ginger Chicken
Carrot Cake

ENERGY NEEDS – Question 2

Jane keeps a record of what she eats each day. Before dinner on that day her total intake of energy had been 7520 kJ.

Jane does **not** want her total energy intake to go **below or above her recommended daily amount** by more than 500 kJ.

Decide whether the special "Fixed Price Menu" will allow Jane to stay within ±500 kJ of her recommended energy needs. Show your work.

Response Coding guide for ENERGY NEEDS Question 1

Full Credit

Code 1: 12120 kilojoules. If no answer is given, check whether the student circled "12120" in the table.

No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Energy Needs*
Question: *Question 1*
Problem type: *Decision making*
Item type: *Closed-constructed response*
Level: *Below Level 1*
PISA scale score: *361*
Item code: *X430Q01*

The scale score for the item indicates that correct performance on it is still significantly beneath the level of performance that would be judged to be indicative of performing at Level 1, that of a *basic problem solver*. A student who answers it correctly does need to understand the decision-making demands of this problem and be able to locate a table entry by making appropriate links among at least three constraining factors (occupation, age and gender). However, correct performance here does not signify that students have developed a full set of skills that allow them to address problems consistently, showing understanding of the issues and factors at the heart of the problems, or consistently provide solutions to even the easiest of problems appropriate for students of their age level.

The second question in the *Energy Needs* unit was more demanding: full credit performance on the item is indicative of performance at Level 3. This item required a consideration of the case of Jane Gibbs, a 19-year-old high jumper. To answer it, students had to juggle the constraints of age, gender, and activity level. They also had to take account of the kilojoules already consumed in determining whether Jane can have the Fixed Price meal. It is these multiple connections that move the problem to a higher level.

Response Coding guide for ENERGY NEEDS Question 2

Full Credit

Code 2: Food from the fixed price menu does not contain enough energy for Jane to keep within 500 kJ of her energy needs. Work needs to show:

- Calculation of the total energy of the fixed price menu: 355+795+565=1715.
- That the difference between 1715+7520 and 9820 is more than 500.

Partial Credit

Code 1: Correct method, but some minor errors in the computation leading to either correct or incorrect, but consistent, conclusion.

Or

Correctly adding up the total energy for the fixed price menu (1715 kJ), but incorrectly interpreting the question.

- 1715 is above 500 kJ, so Jane should not have this.
- 1715+7520=9235. This is within 500 of 8780, so “Yes”.
- Correct calculations, but concludes “Yes” or gives no conclusion.

No Credit

Code 0: Other responses, including "No", without explanation.

- No, Jane should not order from the fixed price menu.

Or

Correct reasoning in words but no figures shown. That is, Code 1 needs to have some supporting figures.

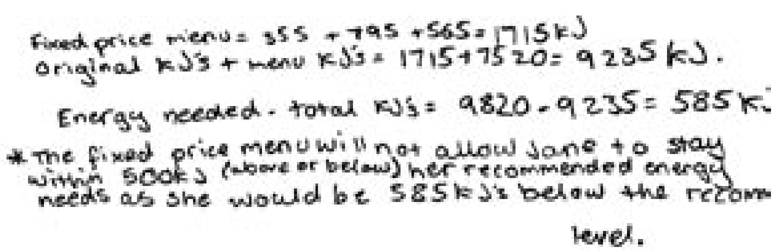
- The fixed price menu does not have enough kJ, so Jane should not have it.

Code 9: Missing.

Unit: *Energy Needs*
Question: *Question 2*
Problem type: *Decision making*
Item type: *Open-constructed response*
Level: *Level 2 (partial credit) and Level 3 (full credit)*
PISA scale score: *587 (partial credit) and 624 (full credit)*
Item code: *X430Q02*

Student work on question 2, such as that shown in Figure 4.2 below, is indicative of Level 3 work. Such work reflects student understanding of the problem of comparing the items on the Fixed Price Menu with their appropriate kilojoule (kJ) ratings in the chart. A comparison of the kJ associated with the Fixed Price dinner, 1715 kJ, and the 7520 kJ that Jane has already consumed gives a total of 9235 kJ. An examination of this data with the kJ intake needs of a female with a heavy activity level and age 19 indicates a need of 9820 kJ per day. Thus, Jane's total only comes to 9235, which is 585 kJ short of the recommended daily level. Hence, the Fixed Price Meal will not allow Jane to meet her recommended Energy Need level, even allowing for the 500 kJ variance mentioned.

Figure 4.2 ■ **Full credit student work on *Energy Needs*, Question 2**



Students at Level 3 are considered *reflective, communicative problem solvers* because of their ability to manage all of the factors involved in a problem, in a clear, correct and coherent way. In addition to their systematic approach to a solution, these students are also capable of communicating the results to others.

Students failing to receive full credit on Question 2 have the opportunity of receiving partial credit. Figure 4.3 shows the work of a student who used a correct method, but lost track of the process after finding the kJ needed and the total in the fixed price menu. This work is indicative of students who lack the ability to reflect on and monitor their work.

Figure 4.3 ■ **Partial credit student work on *Energy Needs,* Question 2 – example 1**

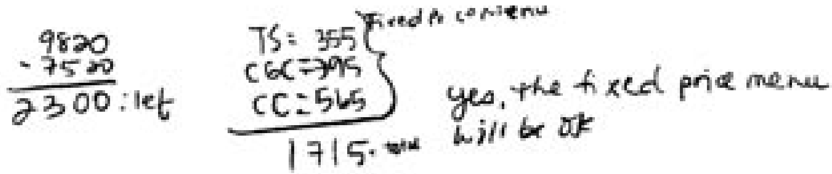
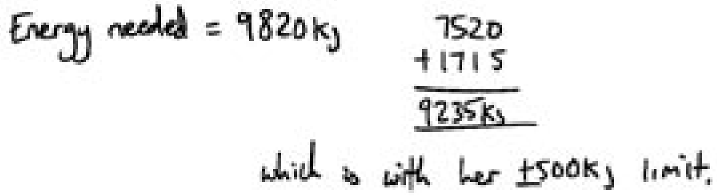


Figure 4.4 shows the work of a student who also received partial credit. This student correctly located the amount needed and correctly added the fixed price menu total to that of the amount already consumed that day. However, the student then drew the wrong conclusion.

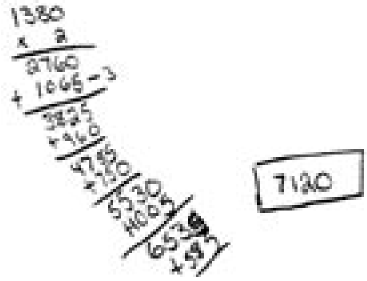
Figure 4.4 ■ **Partial credit student work on *Energy Needs,* Question 2 – example 2**



Partial credit performance associated with Question 2 is located at 587 on the PISA problem-solving scale. This point corresponds with a Level 2 performance. Such work signifies the ability to comprehend a problem, systematically approach it, combine different sources of information when needed and work toward a solution. In the example shown in Figure 4.4, the student found the energy amount needed and carried out the correct calculations, but then made a misstep in making a final judgment.

Finally, there are students who received no credit for their work on Question 2. Figure 4.5 contains a sample of such work. This work reflects a student who worked off-task, in solving another problem – the student apparently tried to figure out how Jane might have achieved a total of 7520 kJ of energy in her other meals that day. Often, however, students who achieved no credit were clearly on the right track but made a combination of arithmetic and logical errors.

Figure 4.5 ■ **No credit student work on *Energy Needs,* Question 2**



CINEMA OUTING

The second *decision-making* unit is *Cinema Outing,* presented as an example of a *decision-making* unit in Chapter 2. This unit centres on the situation of three friends making plans to attend a movie during a week of vacation from school. It is comprised of two questions. Students' responses are based on the general information and the question-specific information contained in the items shown below. *Cinema Outing* as a unit requires the student to read and analyse the information found in the lists of movies and times. The skills involved require careful consideration of relationships, noting the implications determined by the constraints in times, movies previously seen and conflicts in schedules, as well as parental concerns.

This problem is about finding a suitable time and date to go to the cinema.

Isaac, a 15-year-old, wants to organise a cinema outing with two of his friends, who are of the same age, during the one-week school vacation. The vacation begins on Saturday, 24th March and ends on Sunday, 1st April.

Isaac asks his friends for suitable dates and times for the outing. The following information is what he received.

Fred: "I have to stay home on Monday and Wednesday afternoons for music practice between 2:30 and 3:30."

Stanley: "I have to visit my grandmother on Sundays, so it can't be Sundays. I have seen Pokamin and don't want to see it again."

Isaac's parents insist that he only goes to movies suitable for his age and does not walk home. They will fetch the boys home at any time up to 10 p.m.

Isaac checks the movie times for the vacation week. This is the information that he finds.

TIVOLI CINEMA

Advance Booking Number: 01924 423000
24 hour phone number: 01924 420071
Bargain Day Tuesdays: All films $3
Films showing from Fri 23rd March for two weeks:

Children in the Net 113 mins 14:00 (Mon-Fri only) 21:35 (Sat/Sun only)	Suitable only for persons of 12 years and over	**Pokamin** 105 mins 13:40 (Daily) 16:35 (Daily)	Parental Guidance. General viewing, but some scenes may be unsuitable for young children
Monsters from the Deep 164 mins 19:55 (Fri/Sat only)	Suitable only for persons of 18 years and over	**Enigma** 144 mins 15:00 (Mon-Fri only) 18:00 (Sat/Sun only)	Suitable only for persons of 12 years and over
Carnivore 148 mins 18:30 (Daily)	Suitable only for persons of 18 years and over	**King of the Wild** 117 mins 14:35 (Mon-Fri only) 18:50 (Sat/Sun only)	Suitable for persons of all ages

CINEMA OUTING – Question 1

Taking into account the information Isaac found on the movies, and the information he got from his friends, which of the six movies should Isaac and the boys consider watching?

Circle "Yes" or "No" for each movie.

Movie	Should the three boys consider watching the movie?
Children in the Net	Yes / No
Monsters from the Deep	Yes / No
Carnivore	Yes / No
Pokamin	Yes / No
Enigma	Yes / No
King of the Wild	Yes / No

CINEMA OUTING – Question 2

If the three boys decided on going to "Children in the Net", which of the following dates is suitable for them?

A. Monday, 26th March

B. Wednesday, 28th March

C. Friday, 30th March

D. Saturday, 31st March

E. Sunday, 1st April

Unit: *Cinema Outing*
Question: *Question 2*
Problem type: *Decision making*
Item type: *Multiple choice*
Level: *Level 1*
PISA scale score: *468*
Item code: *X601Q02*

The analysis of students' responses to *Cinema Outing*, Question 2 indicated that it was more accessible to students than *Cinema Outing*, Question 1. This may have resulted from the fact that students could answer Question 2 by eliminating films on the basis of a conflict with one constraint. A correct response to *Cinema Outing*, Question 1 required attendance to multiple constraints at one time. This criterion of the number of constraints that must be attended to at one time is a major discriminator between performance at Level 1 and Level 2.

Response Coding guide for CINEMA OUTING Question 2

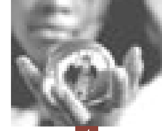
Full Credit

Code 1: C. Friday, 30th March.

No Credit

Code 0: Other responses.

Code 9: Missing.



Unit: *Cinema Outing*
Question: *Question 1*
Problem type: *Decision making*
Item type: *Multiple choice*
Level: *Level 1 (partial credit) and Level 2 (full credit)*
PISA scale score: *442 (partial credit) and 522 (full credit)*
Item code: *X601Q01*

Cinema Outing, Question 1 calls on students to show they understand the constraints imposed by the problem situation and to make a decision about whether there is even a possibility that the three boys can see a given movie together. A correct answer is linked to a score of 522 on the problem-solving scale. Students who correctly respond to this item are able to make decisions about whether a given alternative is possible or not. This indicates understanding and the ability to interpret constraints in multiple combinations with regard to each of the possible alternatives. Such problem-solving performances are judged to be indicative of Level 2 problem-solving work.

Response Coding guide for CINEMA OUTING Question 1

Full Credit
Code 2: Yes, No, No, No, Yes, Yes, in that order.

Partial Credit
Code 1: One incorrect answer.

No Credit
Code 0: Other responses.
Code 9: Missing.

Students who answered all but one of the questions correctly received partial credit for the item. However, such responses were placed lower on the problem-solving scale at Level 1, with 442 points.

HOLIDAY

The third *decision-making* unit, *Holiday,* asks two questions of students. It deals with the planning of a route and places to stay overnight on a holiday trip. Students were presented with a map and a chart showing the distances between the towns illustrated on the map.

This problem is about planning the best route for a holiday.

Figures 1 and 2 show a map of the area and the distances between towns.

Figure 1. Map of roads between towns

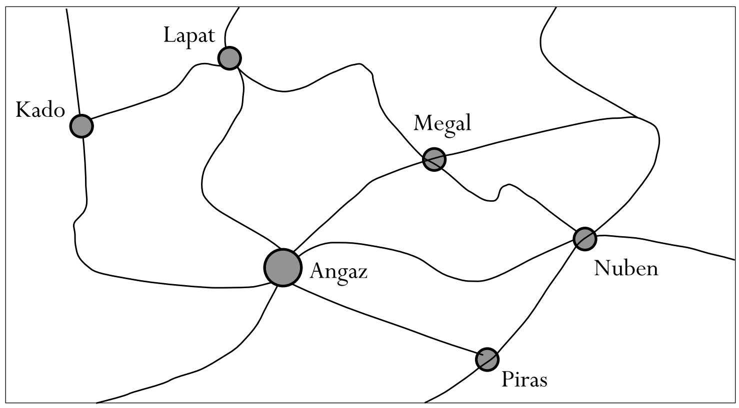


Figure 2. Shortest road distance of towns from each other in kilometres.

	Angaz	Kado	Lapat	Megal	Nuben	Piras
Angaz						
Kado	550					
Lapat	500	300				
Megal	300	850	550			
Nuben	500		1000	450		
Piras	300	850	800	600	250	

HOLIDAY – Question 1

Calculate the shortest distance by road between Nuben and Kado.

Distance: .... kilometres.

HOLIDAY – Question 2

Zoe lives in Angaz. She wants to visit Kado and Lapat. She can only travel **up to 300 kilometres** in any one day, but can break her journey by camping overnight anywhere between towns.

Zoe will stay for two nights in each town, so that she can spend one whole day sightseeing in each town.

Show Zoe's itinerary by completing the following table to indicate where she stays each night.

Day	Overnight Stay
1	Camp-site between Angaz and Kado.
2	
3	
4	
5	
6	
7	Angaz

Unit: *Holiday*
Question: *Question 1*
Problem type: *Decision making*
Item type: *Closed-constructed response*
Level: *Level 2*
PISA scale score: *570*
Item code: *X602Q01*

The stimulus materials for *Holiday,* Question 1 require little in terms of reading text. However, students must read and interpret information from the map and from the distance chart. Some of the distances that they have to find in the chart require them to read distances starting from the bottom of the chart, rather than from the left down. For example, in determining the distance from Nuben to Piras, one needs to transform the search to that of finding the distance from Piras to Nuben. Examining students' work, one sees a variety of interpretations or errors that they made in responding to Question 1. For example, a student who responds 1 100 kilometres may have determined the distance of the route that goes from Nuben to Piras to Angaz to Kado. Other students may have identified the shortest route, Nuben-Angaz-Kado, but calculated the distance incorrectly.

Response Coding guide for HOLIDAY Question 1

Full Credit
Code 1: 1 050 kilometres.

No Credit
Code 0: Other responses.
- Nuben-Angaz-Kado, no distance given.

Code 9: Missing.

The second *Holiday* question was associated with performances scoring higher on the PISA scale. It asked students to make a decision about how to schedule their travel among the towns in terms of overnight stays.

Unit: *Holiday*
Question: *Question 2*
Problem type: *Decision making*
Item type: *Open-constructed response*
Level: *Level 2 (partial credit) and Level 3 (full credit)*
PISA scale score: *593 (partial credit) and 603 (full credit)*
Item code: *X602Q02*

This question set a number of constraints that needed to be complied with simultaneously – a maximum of 300 km travelled in a given day, starting and finishing in Zoe's hometown of Angaz, visiting Kado and Lapat, and staying two nights in each of these cities so that she can achieve her vacation goals. Note that while full credit, associated with Level 3, could only be obtained for a fully correct answer, partial credit could be obtained with one entry incorrect, linked to a score right at the top of Level 2, only 11 score points below a fully correct answer at the bottom of Level 3. A student who made one mistake in calculating the answer to this problem was still able to go through the main steps of analytic reasoning required to solve it.

Response Coding guide for HOLIDAY Question 2

Full Credit

Code 2: Entries as shown below:

Day	Overnight Stay
1	Camp-site between Angaz and Kado.
2	Kado
3	Kado
4	Lapat
5	Lapat
6	Camp-site between Lapat and Angaz (OR just "Camp-site")
7	Angaz

Partial Credit

Code 1: One error. An error means the entry is not correct for the corresponding day.

- "Sightseeing in Lapat" for day 3.
- A town name for day 6.
- No entry for day 6.

No Credit

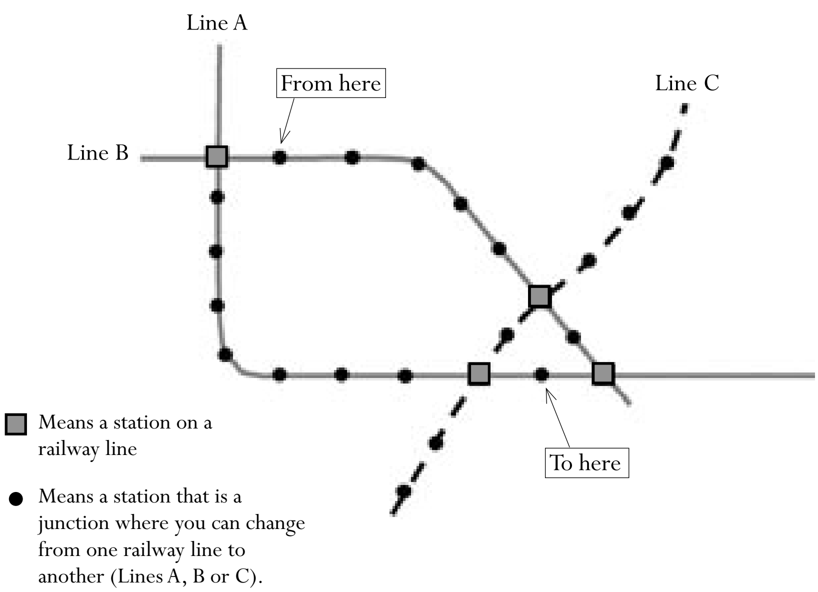
Code 0: Other responses.

Code 9: Missing.

TRANSIT SYSTEM

This single-item unit requires students to look at a transit system map and information about fares and travel times and to calculate the best route, the cost and the time taken to travel between two stations.

The following diagram shows part of the transport system of a city in Zedland, with three railway lines. It shows where you are at present, and where you have to go.



The fare is based on the number of stations travelled (not counting the station where you start your journey). Each station travelled costs 1 zed.

The time taken to travel between two adjacent stations is about 2 minutes.

The time taken to change from one railway line to another at a junction is about 5 minutes.

TRANSIT SYSTEM – Question 1

The diagram indicates a station where you are currently at ("From here"), and the station where you want to go ("To here"). **Mark on the diagram** the best route in terms of cost and time, and indicate below the fare you have to pay, and the approximate time for the journey.

Fare: .... zeds.

Approximate time for journey: .... minutes.

Unit: *Transit System*
Question: *Question 1*
Problem type: *Decision making*
Item type: *Open-constructed response*
Level: *Level 3 (partial and full credit)*
PISA scale score: *608 (partial credit) and 725 (full credit)*
Item code: *X415Q01*

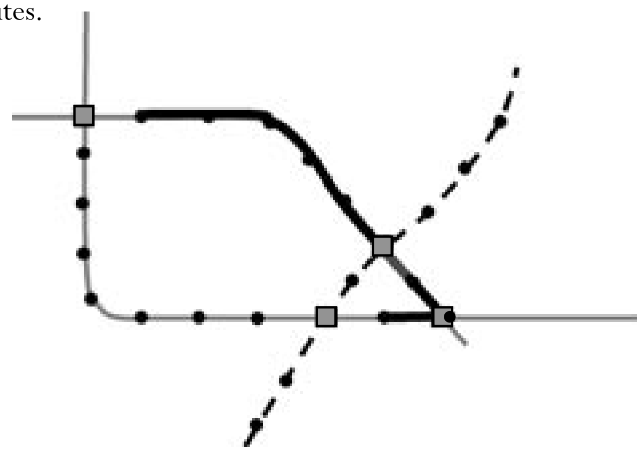
Somewhat surprisingly, this *decision-making* item turned out to be the most difficult problem-solving item on the assessment. This may be due to the amount of external information that it required students to bring to the solution of the problem posed. While there was no indication through different country results that students in some countries might have been more familiar with the use of a transit system, it appears that many students were unable to deal with the task of transferring from one train to another. In addition, there appeared to be confusion with regard to counting the number of segments on the route between the origin and destination.

The coding guide for the problem awards full credit if students show the correct route, the lowest fare and quickest time for the trip. However, if a student did not mark the map, but correctly gave the fare and time, they also are awarded full credit, as no route other than the correct one has that combination of fare and time.

Response Coding guide for TRANSIT SYSTEM Question 1

Full Credit

Code 21: Route as shown: fare 8 zeds; approximate time for journey: 21 minutes.



Code 22: No route shown; fare 8 zeds; time 21 minutes.

Partial Credit

Code 11: Best route shown, with correct fare or time, but not both.
- Best route shown; fare 8 zeds; time 26 minutes.
- Best route shown; fare missing; time 21 minutes.

Code 12: One of the two other possible routes shown, with correct fare and time for that route.
- Route shown is the one that first goes left; fare 10 zeds; time 25 minutes.
- Route shown is the one via Lines B, C & A; fare 8 zeds; time 26 minutes.

Code 13: No route shown, but correct fare AND time for one of the two other routes are given.

- No route shown; fare 10 zeds; time 25 minutes.
- No route shown; fare 8 zeds; time 26 minutes.

No Credit

Code 01: Best route shown, but both fare and time incorrect or missing.

- Best route shown; fare missing; time 26 minutes.

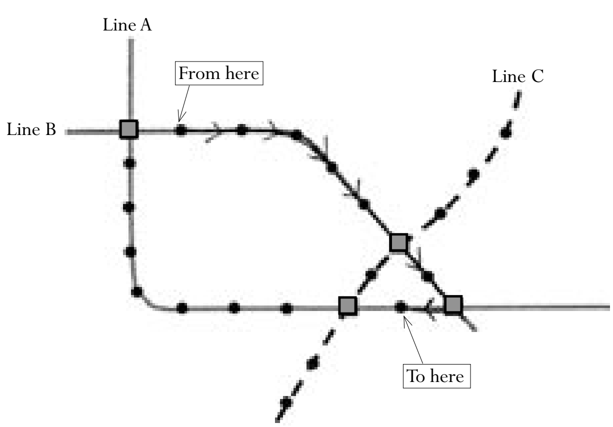
Code 02: Other responses.

- Lines B, C & A route shown; fare and time missing.

Code 99: Missing. (Note that Code 99 should only be given when there is no route shown AND no fare given AND no time given.)

Note that the coding guide makes use of double-digit coding, which allows for the differentiation of students' responses in the data interpretation. Double-digit coding uses the first digit of the response code to indicate the score associated with a student's work. The second digit of the response code either indicates the method the student used in gaining that coding, or the type of error the student made. For example, a piece of student work that was assigned partial credit code 11 is shown in Figure 4.6. Here the student indicated the correct path and fare, but an incorrect journey time.

Figure 4.6 ■ **Partial credit solution for *Transit System*** (Response Coding Code 11)



Fare:........8........ zeds.

Approximate time for journey:........19........ minutes.

This coding makes it possible to analyse what it was that made students find this problem so difficult. Many found it hard to factor in the transfer time when changing from one line to another and apparently had difficulty dealing with the fact that the transfer time is in addition to the travel time. Another error source commonly observed in students' work was that they incorrectly counted stations or the links between stations in calculating costs and time, respectively.

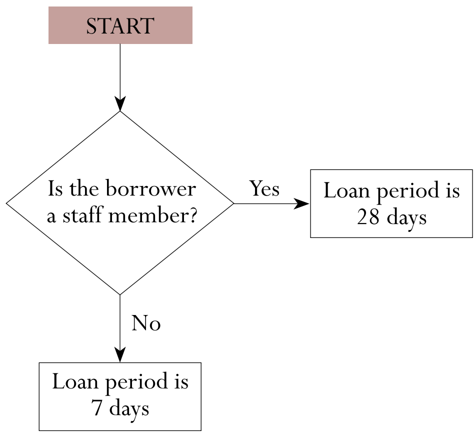
System analysis and design units

PISA included four units assessing students' capabilities to solve problems involving *system analysis and design*. One of these units had three items, two had two items each and the fourth unit had one item. *System analysis and design* problems differ from the *decision-making* items in that not all of the possible options are given nor are the constraints as obvious. In the *system analysis and design* problems, students have to develop an understanding of the problem, beginning with the identification of the relationships existing between the parts of the system, or to design a system with certain relationships among its main features. Next, students have to develop a representation that brings the inherent relationships into a manipulative form. From here the students can test the system or design by working with individual or sets of related features in the system. Finally, students are generally involved in justifying their analysis or defending their design.

LIBRARY SYSTEM

The *system analysis and design* unit with the most accessible question was the *Library System* unit. This unit contained two items. The first question required students to interpret rules for a system, identify which of the rules applied, and apply them. A correct response indicated that they understood the general nature of such a system. The second question involved the students in developing a flow chart that would implement a given set of rules.

The **John Hobson High School** library has a simple system for lending books: for staff members the loan period is 28 days and for students the loan period is 7 days. The following is a decision tree diagram showing this simple system:



The **Greenwood High School** library has a similar, but more complicated, lending system:

- All publications classified as "Reserved" have a loan period of 2 days.
- For books (not including magazines) that are **not** on the reserved list, the loan period is 28 days for staff, and 14 days for students.
- For magazines that are **not** on the reserved list, the loan period is 7 days for everyone.
- Persons with any overdue items are not allowed to borrow anything.

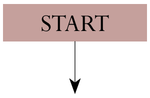
LIBRARY SYSTEM – Question 1

You are a student at **Greenwood High School**, and you do not have any overdue items from the library. You want to borrow a book that is **not** on the reserved list. How long can you borrow the book for?

Answer: .... days.

LIBRARY SYSTEM – Question 2

Develop a decision tree diagram for the **Greenwood High School Library** system so that an automated checking system can be designed to deal with book and magazine loans at the library. Your checking system should be as efficient as possible (*i.e.* it should have the least number of checking steps). Note that each checking step should have only **two** outcomes and the outcomes should be labelled appropriately (e.g. "Yes" and "No").



Unit: *Library System*
Question: *Question 1*
Problem type: *System analysis and design*
Item type: *Closed-constructed response*
Level: *Level 1*
PISA scale score: *437*
Item code: *X402Q01*

Using the rules of the system, the first question required students to determine for how long a student could borrow a book that was not on the reserve list. To respond correctly, the student had to understand the rules, recognize which ones applied to non-reserve book loans to students, and then determine the loan period.

Response Coding guide for LIBRARY SYSTEM Question 1

Full Credit

Code 1: 14 days.

No Credit

Code 0: Other responses.

Code 9: Missing.

This question is located at Level 1 because it only requires that students understand the nature of a problem and locate and retrieve information related to a major feature of the problem. In this case the check-out policies involved were well-defined and fairly easy to implement by checking the conditions given.

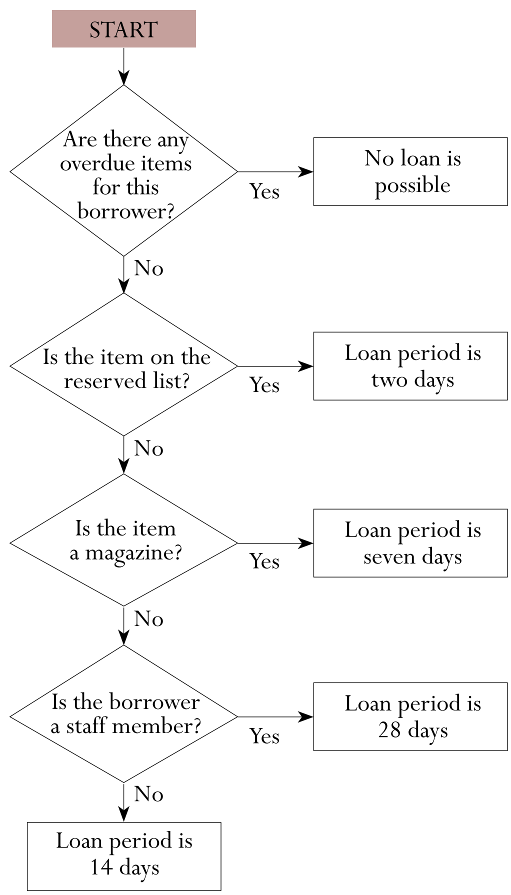
Unit: *Library System*
Question: *Question 2*
Problem type: *System analysis and design*
Item type: *Open-constructed response*
Level: *Level 3 (partial and full credit)*
PISA scale score: *658 (partial credit 1), 677 (partial credit 2) and 693 (full credit)*
Item code: *X402Q02*

The second library question was more difficult. It asked students to develop a flow chart illustrating the set of rules applied from a written list. The chart was to provide a design for developing an automated system that would provide the loan period for a given library item. Student responses could receive full credit, one of two partial credit scores according to the criteria listed overleaf, or no credit. The analysis of student work indicated that both full and partial credit responses are indicative of Level 3 performance. To construct the flow chart, even with minor errors, students must not only develop an understanding of the multiple regulations and the relationships between them for this library system, but organise and monitor their approach to constructing and communicating the solution. Handling such a problem in its totality requires students to keep track of the interrelationships involved even for the partial credit responses.

Response Coding guide for LIBRARY SYSTEM Question 2

Full Credit

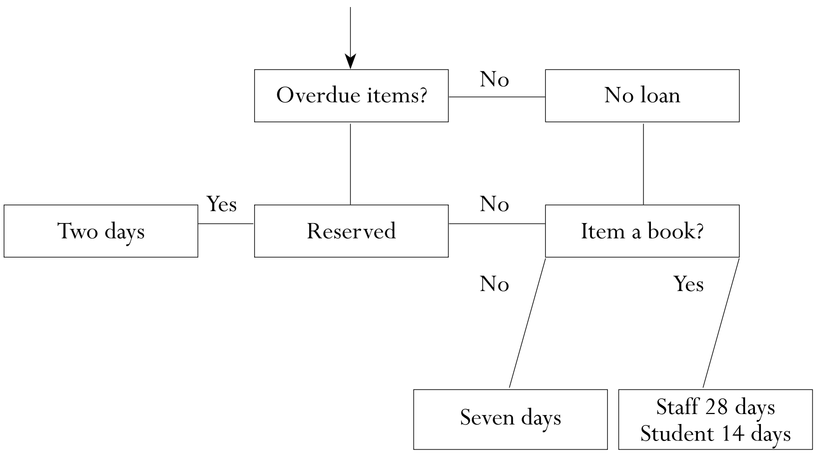
Code 31: The most efficient system is a four-step check system as follows:



Partial Credit

Code 21: The four check steps are in the right sequence, but there is a minor error. For example:

- One loan period is incorrect.
- One loan period is missing.
- One or more Yes/No missing.
- One Yes/No incorrectly labelled. For example:



Code 22: The check for "overdue items" is written as a statement outside the decision tree diagram, but the other three check steps are completely correct and in the right sequence.

Code 23: Two check steps are out of order, resulting in five steps, as one extra check step is required. The system is still complete, but less efficient. Complete means that the checking system will produce the correct loan periods in all cases.

Code 11: The diagram is correct except that the first three check steps are out of order in one (but not both) of the following two ways:

- The checks for "reserved list" and "magazine" are interchanged.
- The checks for "overdue items" and "reserved list" are interchanged.

Code 12: The check for "overdue items" is written as a statement outside the decision tree diagram. The other three check steps are in the right sequence, but with a minor error.

OR

The check for "overdue items" is missing, but the other three check steps are completely correct and in the right sequence.

No Credit

Code 01: The system is complete, but has more than five check steps.

Code 02: Other responses.

- System incomplete and is not covered by any of the partial credit codes.
- Five or more check steps, and the system is incomplete.
- Five check steps, with "overdue items" missing.
- A checking step has more than two outcomes.

Code 99: Missing.

Figure 4.7 provides an example of student work receiving full credit. This student has presented the work in a clear and concise fashion and followed the instructions that each decision point should have exactly two outcomes.

Figure 4.7 ■ **Example of full credit response to *Library System*, Question 2**

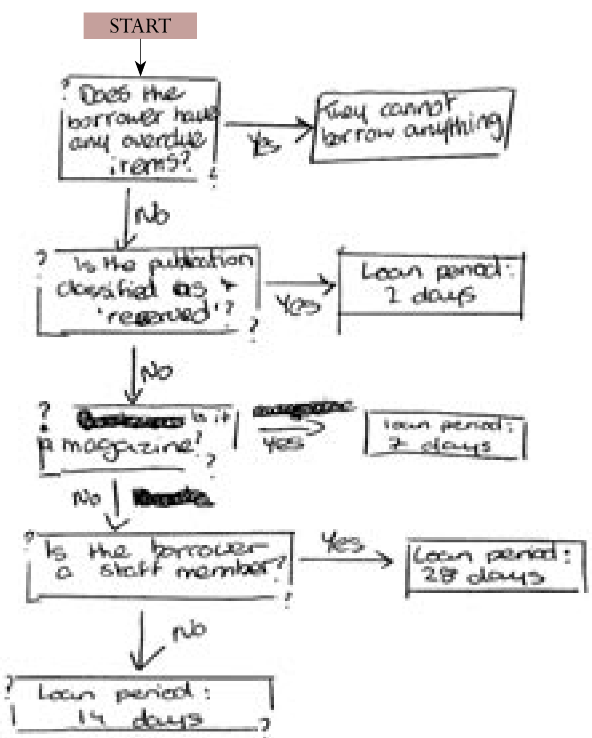
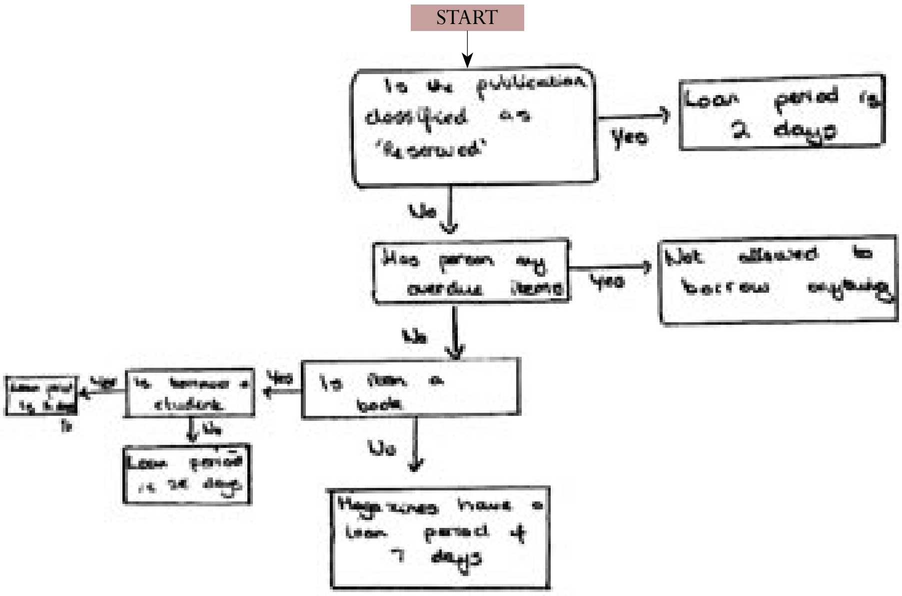


Figure 4.8 shows an example of student work receiving a partial credit code of 11. This double-digit partial credit score indicates that the student's work is correct except that the first three check steps are out of order in one (but not both) of the following two ways:

- The checks for "reserved list" and "magazine" are interchanged.
- The checks for "overdue items" and "reserved list" are interchanged.

Figure 4.8 ■ **Partial credit solution for *Library System,* Question 2** (Response Code 11)



In fact, an examination of the rules and the student's work shows that this student has reversed the order of the checks for overdue books and items listed on the reserved list, but is otherwise correct. The overdue books criterion should have been listed first before any consideration is given to the length of any checkout period. Most students had difficulty in correctly handling the overdue books criterion.

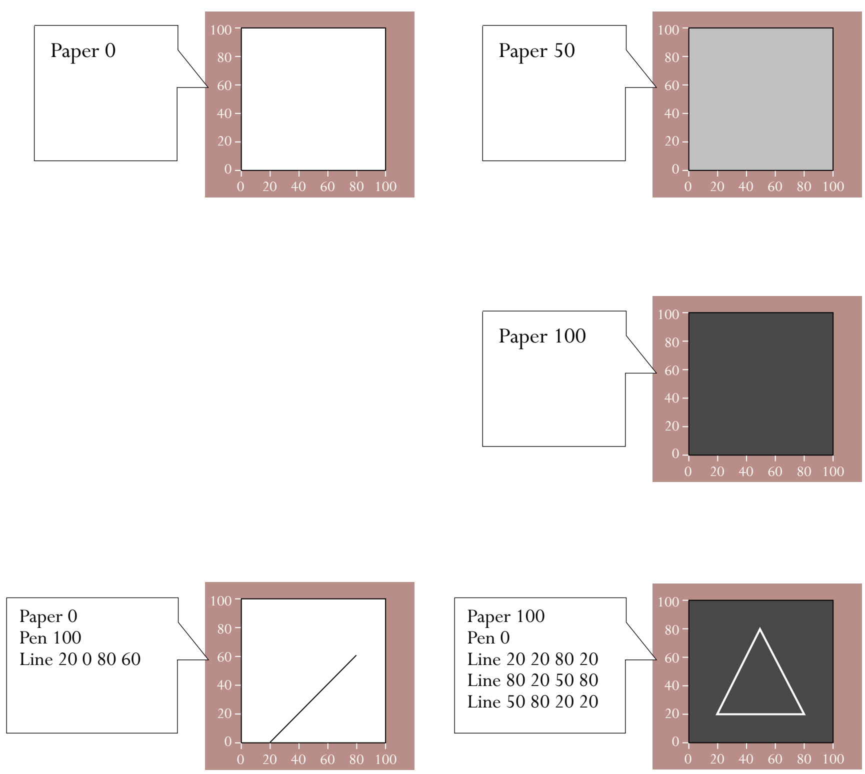
The *Library System* unit was interesting in that it contained two questions of contrasting difficulty. The first question was the second easiest problem-solving item in PISA, while the second question was the second hardest overall. Thus in the framework of a single situation, problem-solving abilities at very different levels of difficulty could be assessed.

DESIGN BY NUMBERS©

The *system analysis and design* unit containing the next easiest question was the *Design by Numbers©* unit. This unit had three questions, the most of any unit in the assessment. The *Design by Numbers©* unit was based on a context involving a language for graphical design developed by the Aesthetics and Computation Group at the MIT Media Laboratory. The system at the heart of this problem is a programming language that relates written instructions to actions taking place on a monitor-like screen on the page.

Design by Numbers is a design tool for generating graphics on computers. Pictures can be generated by giving a set of commands to the program.

Study carefully the following example commands and pictures before answering the questions.



1. Design by Numbers was developed by the Aesthetics and Computation Group at the MIT Media Laboratory. Copyright 1999, Massachusetts Institute of Technology. The program can be downloaded from *http://dbn.media.mit.edu*.

DESIGN BY NUMBERS© – Question 1

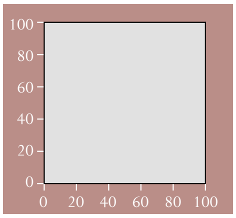
Which of the following commands generated the graphic shown below?

A. Paper 0

B. Paper 20

C. Paper 50

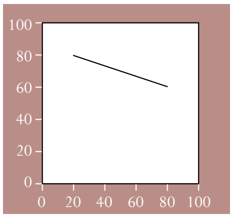
D. Paper 75



DESIGN BY NUMBERS© – Question 2

Which of the following set of commands generated the graphic shown below?

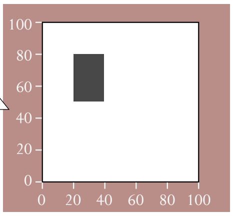
A.	Paper 100	Pen 0	Line 80 20 80 60
B.	Paper 0	Pen 100	Line 80 20 60 80
C.	Paper 100	Pen 0	Line 20 80 80 60
D.	Paper 0	Pen 100	Line 20 80 80 60



DESIGN BY NUMBERS© – Question 3

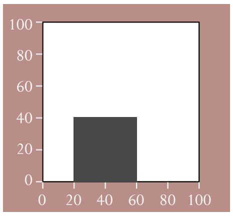
The following shows an example of the "Repeat" command.

```
Paper 0
Pen 100
Repeat A 50 80
{
Line 20 A 40 A
}
```



The command "Repeat A 50 80" tells the program to repeat the actions in brackets { }, for successive values of A from A=50 to A=80.

Write commands to generate the following graphic:



Unit: *Design by Numbers©*
Question: *Question 1*
Problem type: *System analysis and design*
Item type: *Multiple choice*
Level: *Level 2*
PISA scale score: *544*
Item code: *X412Q01*

The first question assessed students' ability to analyse a series of examples and abstract out the relationship between the shade level of the screen and the programming command line associated with that outcome. In this case, the sample screens show the commands of **Paper**, **Pen**, and **Line**. Comparing and contrasting the examples present, an analysis indicates that the command **Paper** is the appropriate command to give a blank shaded screen. In order to select **Paper** 20 as the appropriate answer, the student must compare the "paper" shadings shown.

Response Coding guide for DESIGN BY NUMBERS© Question 1

Full Credit

Code 1: B. Paper 20.

No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Design by Numbers©*
Question: *Question 2*
Problem type: *System analysis and design*
Item type: *Multiple choice*
Level: *Level 2*
PISA scale score: *553*
Item code: *X412Q02*

The second question in the *Design by Numbers©* unit involves a similar task to the first. However, the analysis required here involves the abstraction of a sequenced set of commands rather than a single command. The solution can still be found by selecting one of the alternatives in a multiple choice setting, but students have to distinguish between the effects of the values associated with Paper and Pen commands and then understand the coordinate structure for the Line command.

Response Coding guide for DESIGN BY NUMBERS© Question 2

Full Credit

Code 1: D. Paper 0 Pen 100 Line 20 80 80 60.

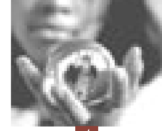
No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Design by Numbers©*
Question: *Question 3*
Problem type: *System analysis and design*
Item type: *Open-constructed response*
Level: *Level 2 (partial credit) and Level 3 (full credit)*
PISA scale score: *571 (partial credit) or 600 (full credit)*
Item code: *X412Q03*

The final item in the unit involves the design of a set of commands that will result in the design tool replicating a given figure on the screen. Here the student has to bring together the effects associated with the previous commands and add to that the Repeat command. This involves realising that the figure can be formed by drawing lines repeatedly, creating an appropriate command that needs to be repeated and then establishing the values that the program has to cycle through in the repetition pattern to allow embedded command lines to draw the figure. This involves more analysis than the previous items in the unit and has the added design feature of students having to write the command lines, rather than select them.



Response Coding guide for DESIGN BY NUMBERS© Question 3

Full Credit

Code 2: Correct commands.

- Note that in the Repeat command, 0 and 40 can be switched (*i.e.* Repeat 40 0). In the command Line 20 A 60 A, 20 and 60 can be switched (*i.e.* Line 60 A 20 A).

```
Paper 0
Pen 100
Repeat A 0 40
{
Line 20 A 60 A
}
```

- Note that in the Repeat command, 20 and 60 can be switched (*i.e.* Repeat 60 20). In the command Line A 0 A 40, 0 and 40 can be switched (*i.e.* Line A 40 A 0).

```
Paper 0
Pen 100
Repeat A 20 60
{
Line A 0 A 40
}
```

(In short, 0 and 40 should be in the Y position and 20 and 60 should be in the X position.)

Partial Credit

Code 1: Correct commands but incorrect placement of numbers in the Line command.

- ```
  Paper 0
  Pen 100
  Repeat A 20 60
  {
  Line 0 A 40 A
  }
  ```

Correct commands, but one incorrect number in either the Repeat or the Line command. Note that if there is any number other than 0 or 20 or 40 or 60 (*e.g.* 50 or 80 are used), or if the same number is repeated in the one command, then Code 0 should be given.

- ```
  Pen 100
  Paper 0
  Repeat A 0 40
  {
    Line 0 A 60 A
  }
  ```

Correct Repeat section, but missing or incorrect Paper or Pen command.

- Repeat y 0 40
  {
  Line 20 y 60 y
  }

Correct numbers, but a small mistake in the Line command.

- Paper 0
  Pen 100
  Repeat A 20 60
  {
  A 0 A 40
  }

No Credit

Code 0: Other responses.

- Paper 0
  Pen 100
  Line 20 0 60 40

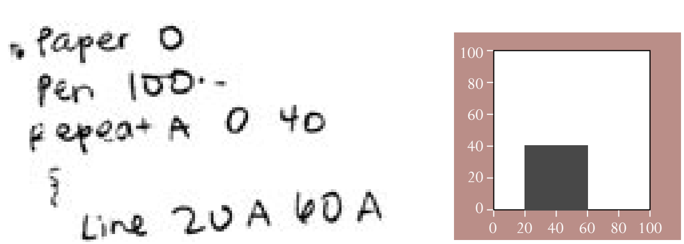
- Paper 0
  Pen 100
  Repeat A 20 60
  {
  Line A 20 A 60
  }

Code 9: Missing.

The work in Figure 4.9 shows the performance of a student who responds to the requirements of question three. The demands of this item place it in Level 3. (Note that the student's omission of the final } is not considered a significant error in this context. As the response coding guide of Question 3 indicates, minor syntax errors were disregarded in coding students' solution.)

Figure 4.9 ■ **Example of full credit response for *Design by Numbers©*, Question 3**

Write commands to generate the following graphic:



In designing such a program, a student considers and sequences a number of conditions, and determines the commands and variable values that will produce the desired result. Then the student has to write the command line and think through the potential output in a sequential manner, keeping in mind the impact of command on the final output. The mental load and the complexity required in developing the required design as well as the written communication of the design program places this item at Level 3.

Students who made a significant error in one of the command lines in terms of an incorrect value or missed one of the Paper or Pen commands were awarded partial credit. Such performances were considered Level 2 work and were placed on the PISA problem-solving scale at 571.

COURSE DESIGN

This third *system analysis and design* unit contained one open-constructed response item. The item involves sequencing 12 courses of study over a three-year period when some courses have to be taken prior to other courses. It is similar to the normal planning of an individual's academic schedule when there are a number of prerequisite relationships among the courses.

A technical college offers the following 12 subjects for a three-year course, where the length of each subject is one year:

	Subject Code	Subject Name
1	M1	Mechanics Level 1
2	M2	Mechanics Level 2
3	E1	Electronics Level 1
4	E2	Electronics Level 2
5	B1	Business Studies Level 1
6	B2	Business Studies Level 2
7	B3	Business Studies Level 3
8	C1	Computer Systems Level 1
9	C2	Computer Systems Level 2
10	C3	Computer Systems Level 3
11	T1	Technology and Information Management Level 1
12	T2	Technology and Information Management Level 2

COURSE DESIGN – Question 1

Each student will take four subjects per year, thus completing 12 subjects in three years.

A student can only take a subject at a higher level if the student has completed the lower level(s) of the same subject in a previous year. For example, you can only take Business Studies Level 3 after completing Business Studies Levels 1 and 2.

In addition, Electronics Level 1 can only be taken after completing Mechanics Level 1, and Electronics Level 2 can only be taken after completing Mechanics Level 2.

Decide which subjects should be offered for which year, by completing the following table. Write the subject codes in the table.

	Subject 1	Subject 2	Subject 3	Subject 4
Year 1				
Year 2				
Year 3				

Unit: *Course Design*
Question: *Question 1*
Problem type: *System analysis and design*
Item type: *Open-constructed response*
Level: *Level 3 (partial credit and full credit)*
PISA scale score: *602 (partial credit) and 629 (full credit)*
Item code: *X414Q01*

This problem is a different type of *system analysis and design* problem from the two units already discussed. It involves combinatorial reasoning, compared to the analytical reasoning that has played a central role in the previous problems. In *Course Design,* students have to examine carefully the given relationships among the courses. Strong students note the central role played by course sequences of three years, slotting them into the school programme first. Then they place the courses lasting two years and finally one year. In doing so, students must also keep in mind a cross-course sequence requirement that Mechanics 1 (2) must be taken before Electronics 1 (2).

Student performance on the *Course Design* task identified two levels of response. Full credit responses linked to the PISA problem-solving scale at 629, while the partial credit responses were placed at 602. Both of these values indicate student performances at Level 3. This level is appropriate given the number of interacting, interrelated courses in the problem. Further, these courses must be manipulated and placed in the schedule at the same time, as one placement creates or closes possibilities for other placements with each shift of a course.

Response Coding guide for COURSE DESIGN Question 1

Full Credit

Code 2: The order of subjects within a year is unimportant, but the list of subjects for each year should be as given below:

	Subject 1	Subject 2	Subject 3	Subject 4
Year 1	B1	M1	T1	C1
Year 2	B2	M2	E1	C2
Year 3	B3	T2	E2	C3

Partial Credit

Code 1: Mechanics does not precede Electronics. All other constraints are satisfied.

No Credit

Code 0: Other responses.

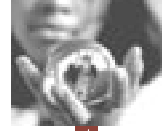
- Table completely correct except that E2 is missing and E1 is repeated where E2 should be or this cell is empty.

Code 9: Missing.

The coding guide shown in the response coding guide of Question 1 gives an example of what a correct response to the item would be. A response receiving partial credit is shown in Figure 4.10. Partial credit responses are characterised by having the courses with sequence numbers correctly placed in the schedule, but without satisfying the requirement that Mechanics must precede Electronics.

Figure 4.10 ■ **Example of partial credit response for *Course Design*, Question 1**

	Module 1	**Module 2**	**Module 3**	**Module 4**
Year 1	B1	C1	M1	E2
Year 2	B2	C2	E1	T1
Year 3	B3	C3	M2	T2



CHILDREN'S CAMP

The final *system analysis and design* problem also involves combinatorial reasoning. This item contains a common system problem of assignment of classes of people to positions consistent with specified relationships between the classes and between the people within the classes. These relationships concern adult-child, male-female and the dormitory size. The manipulation of these is made somewhat more difficult by the different dormitory sizes and the fact that there are eight adults and seven dormitories – hence one dormitory will have two adults.

The Zedish Community Service is organising a five-day Children's Camp. 46 children (26 girls and 20 boys) have signed up for the camp, and 8 adults (4 men and 4 women) have volunteered to attend and organise the camp.

Table 1. **Adults**

Mrs Madison
Mrs Carroll
Ms Grace
Ms Kelly
Mr Stevens
Mr Neill
Mr Williams
Mr Peters

Table 2. **Dormitories**

Name	Number of beds
Red	12
Blue	8
Green	8
Purple	8
Orange	8
Yellow	6
White	6

Dormitory rules:

1. Boys and girls must sleep in separate dormitories.
2. At least one adult must sleep in each dormitory.
3. The adult(s) in a dormitory must be of the same gender as the children.

CHILDREN'S CAMP – Question 1

Dormitory Allocation

Fill the table to allocate the 46 children and 8 adults to dormitories, keeping to all the rules.

Name	Number of boys	Number of girls	Name(s) of adult(s)
Red			
Blue			
Green			
Purple			
Orange			
Yellow			
White			

Unit: *Children's Camp*
Question: *Question 1*
Problem type: *System analysis and design*
Item type: *Open-constructed response*
Level: *Level 2 (partial credit) and Level 3 (full credit)*
PISA scale score: *529 (partial credit) and 650 (full credit)*
Item code: *X417Q01*

Student performances on *Children's Camp* indicated that two levels of performance could be distinguished. A full credit response was placed on the PISA problem-solving scale at 650 and a partial response at 529. The full credit response was within the scale interval associated with Level 3 problem-solving performance. This was quite appropriate given the number of interrelated variables and relationships. A partial score response was associated with Level 2 of problem solving. The response coding guidelines for partial credit allows students to violate one or two of the conditions required for a full credit response to the item. As such, it is considerably less stringent in its demands on the problem solver, although it still requires the student to attend to the variables and carry out considerable combinatorial reasoning. However, the load in information processing and relational checking is smaller.

Response Coding guide for CHILDREN'S CAMP Question 1

Full Credit

Code 2: 6 conditions to be satisfied

- Total girls = 26
- Total boys = 20
- Total adults = four female and four male
- Total (children and adults) per dormitory is within the limit for each dormitory
- People in each dormitory are of the same gender
- At least one adult must sleep in each dormitory to which children have been allocated

Partial Credit

Code 1: One or two conditions (mentioned in Code 2) violated. Violating the same condition more than once will be counted as one violation only.

- Forgetting to count the adults in the tally of the number of people in each dormitory.
- The number of girls and the number of boys are interchanged (number of girls = 20, number of boys = 26), but everything else is correct. (Note that this counts as two conditions violated.)
- The correct number of adults in each dormitory is given, but not their names or gender. (Note that this violates both condition 3 and condition 5.)

No Credit

Code 0: Other responses.

Code 9: Missing.

Figures 4.11 and 4.12 contain a full and partial credit response, respectively. Note that in the partial credit response, the student has answered everything correctly with the exception of placing two girls in the White dormitory. The most common mistake leading to a code 1 response was to omit the adults from the count of people in each dormitory.

Figure 4.11 ■ **Example of full credit response for *Children's Camp*, Question 1**

Name	Number of boys	Number of girls	Names(s) of adult(s)
Red	8	x	Mr Neill, Mr Stevens
Blue	x	7	Mrs Bridges
Green	x	7	Mrs Carroll
Purple	x	7	Ms Grace
Orange	7	x	Mr Williams
Yellow	x	5	Ms Kelly
White	5	x	Mr Peters
	=20 boys	=26girls	

(2 spare beds) ⇐ [beside Red]

Figure 4.12 ■ **Example of partial credit response for *Children's Camp*, Question 1**

Name	Number of boys	Number of girls	Names(s) of adult(s)
Red	6	10	Mrs Bridges Ms Carroll.
Blue	0	7	Ms Grace
Green	0	7	Ms Kelly
Purple	7	0	Mr Stevens
Orange	7	0.	Mr Neill
Yellow	5	0	Mr Williams.
White	1	2.	Mr Peters.

The partial credit response to *Children's Camp* shows the important role that re-evaluation and checking play in correctly responding to either a *system analysis* or a *system design* problem. All of the interrelated aspects of the system must be satisfied.

Trouble-shooting units

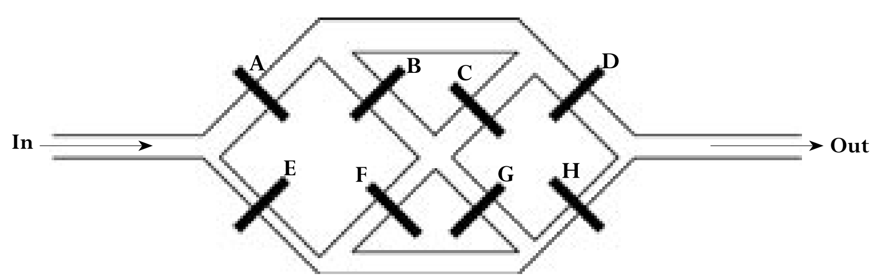
The final two units comprising the PISA problem-solving assessment were drawn from the area of *trouble shooting*. These units have a total of five questions between them. The first problem has three questions and the second has two questions. *Trouble-shooting* units assess students' actions when confronted with a system or mechanism that is underperforming in some way. It may be a non-functioning appliance, such as a videotape recorder, or an appliance such as a sewing machine that just needs an adjustment to correct its performance.

To solve such problems, the student must be able to understand the main features of the system and the actions or responses that are expected of each of these features. Based on this understanding, the student must then be able to identify the causal-response relationships between interrelated parts and the role that such links play in the overall function of the mechanism or system of interest. The student can then diagnose the potential source of the present problem and propose and implement a potential remedy for the problem. Either action should then be evaluated or checked based on the available information for reasonableness or effective repair. Finally, students may need to communicate their solution in writing or through a diagram to explain their thinking and their recommended course of action. Such problems are complicated by the number of interrelated variables involved and the varied number of representations and translations that one might have to make in understanding the system or mechanism from directions or instruction booklets.

IRRIGATION

Irrigation was presented in Chapter 2 as an example of a *trouble-shooting* unit. It contains three questions and involves students in diagnosing the malfunctioning of an irrigation system designed to distribute water to crops on a farm. The system consists of a set of eight gates regulating flow along a system of canals. Each of the gates can either be open or closed. When a gate is closed, no water can flow through it.

Below is a diagram of a system of irrigation channels for watering sections of crops. The gates A to H can be opened and closed to let the water go where it is needed. When a gate is closed no water can pass through it.



This is a problem about finding a gate which is stuck closed, preventing water from flowing through the system of channels.

Michael notices that the water is not always going where it is supposed to.

He thinks that one of the gates is stuck closed, so that when it is switched to open, it does not open.

IRRIGATION – Question 1

Michael uses the settings given in Table 1 to test the gates.

Table 1. **Gate Settings**

A	B	C	D	E	F	G	H
Open	Closed	Open	Open	Closed	Open	Closed	Open

With the gate settings as given in Table 1, **on the diagram below** draw all the possible paths for the flow of water. Assume that all gates are working according to the settings.

IRRIGATION – Question 2

Michael finds that, when the gates have the Table 1 settings, no water flows through, indicating that at least one of the gates set to open is stuck closed.

Decide for each problem case below whether the water will flow through all the way. Circle "Yes" or "No" in each case.

Problem Case	Will water flow through all the way?
Gate **A** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No
Gate **D** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No
Gate **F** is stuck closed. All other gates are working properly as set in Table 1.	Yes / No

IRRIGATION – Question 3

Michael wants to be able to test whether **gate D** is stuck closed.

In the following table, show settings for the gates to test whether **gate D** is stuck closed when it is set to open.

Settings for gates (each one open or closed)

A	B	C	D	E	F	G	H

Unit: *Irrigation* **Question:** *Question 1* **Problem type:** *Trouble shooting* **Item type:** *Open-constructed response* **Level:** *Level 1* **PISA scale score:** *497* **Item code:** *X603Q01*

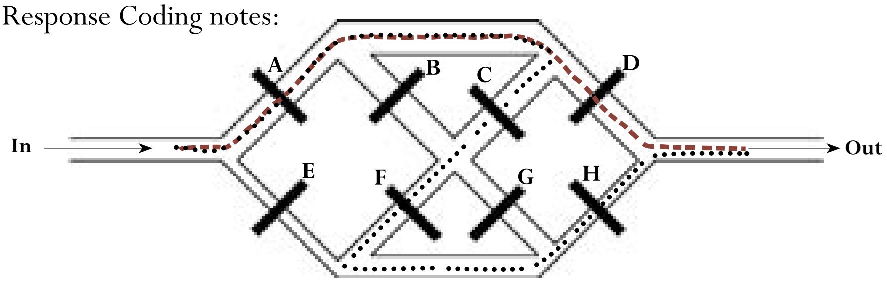
The first question poses an investigation for the students. The results show their understanding of the system of gates and the way(s) in which water flows through the system. The response mode for this question is different from that for other problems in the set. Students have to draw (on the figure provided) the possible paths for water to flow through the system. A correct response is an indication that students are able to note the main features of the system, the gates and the canals, and their states, open/flowing or closed/not flowing. This is the information needed to begin to trouble shoot the system in the following problems.

Response Coding guide for IRRIGATION Question 1

Full Credit

Code 1: Flow paths as shown below:

Response Coding notes:



Ignore any indications of the directions of flow.

Note that the response could be shown IN THE DIAGRAM PROVIDED, OR IN FIGURE 1, OR IN WORDS, OR WITH ARROWS.

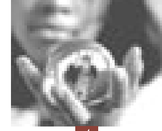
No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Irrigation* **Question:** *Question 2* **Problem type:** *Trouble shooting* **Item type:** *Multiple choice* **Level:** *Level 2* **PISA scale score:** *544* **Item code:** *X603Q02*

The second question in the unit moves to a new level in trouble shooting by further examining students' understanding of the interrelationships among the gates and water flow. Here students are given a complex multiple-choice item with three parts, focusing on what happens when the gate settings are as in Question 1, except that (in each case) one of the gates that is supposed to be open is closed. This raises the problem-solving demands of the problem and shifts to Level 2, as students have to apply reasoning to analyse each of the gate settings and make a decision about whether water will flow all of the way through the system. A correct response requires students to examine the system anew and separate each analysis from the prior and subsequent analyses.



Response Coding guide for IRRIGATION Question 2

Full Credit

Code 1: No, Yes, Yes, in that order.

No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Irrigation*
Question: *Question 3*
Problem type: *Trouble shooting*
Item type: *Open-constructed response*
Level: *Level 2*
PISA scale score: *532*
Item code: *X603Q03*

The third *irrigation* question requires students to develop a test to determine whether gate D is stuck closed. To answer the question correctly, the student must decide on appropriate settings for the gates, knowing the inflow and outflow states of the irrigation system. This is a Level 2 question, as it is slightly easier than Question 2. While the overall item demands are greater in Question 3, Question 2 involves multiple parts which must all be answered correctly to get the item correct.

Response Coding guide for IRRIGATION Question 3

Full Credit

Code 1: A and E are not both closed. D must be open. H can only be open if water cannot get to it (*e.g.* other gates are closed preventing water from reaching H). Otherwise H must be closed.

- H closed, all other gates open

No Credit

Code 0: Other responses.

Code 9: Missing.

FREEZER

The second *trouble-shooting* unit has two questions. This unit deals with diagnosing a probable cause for a malfunctioning home freezer unit. Students confronting this situation have to operate as the freezer user would in the situations described. Information is given in a manual and feedback from the mechanism, in this case the freezer, comes from observing a warning light, the state of the temperature control, and external indications that power is reaching the freezer motor. This problem is clearly one of diagnosing probable causes of a malfunctioning mechanism and hence is a classic *trouble-shooting* problem.

Jane bought a new cabinet-type freezer. The manual gave the following instructions:

- Connect the appliance to the power and switch the appliance on.
- You will hear the motor running now.
- A red warning light (LED) on the display will light up.
- Turn the temperature control to the desired position. Position 2 is normal.

Position	Temperature
1	–15°C
2	–18°C
3	–21°C
4	–25°C
5	–32°C

- The red warning light will stay on until the freezer temperature is low enough. This will take 1 - 3 hours, depending on the temperature you set.
- Load the freezer with food after four hours.

Jane followed these instructions, but she set the temperature control to position 4. After four hours, she loaded the freezer with food.

After eight hours, the red warning light was still on, although the motor was running and it felt cold in the freezer.

FREEZER – QUESTION 2

Jane wondered whether the warning light was functioning properly. Which of the following actions and observations would suggest that the light was working properly?

Circle "Yes" or "No" for each of the three cases.

Action and Observation	Does the observation suggest that the warning light was working properly?
She put the control to position 5 and the red light went off.	Yes / No
She put the control to position 1 and the red light went off.	Yes / No
She put the control to position 1 and the red light stayed on.	Yes / No

FREEZER – Question 1

Jane read the manual again to see if she had done something wrong. She found the following six warnings:

1. *Do not connect the appliance to an unearthed power point.*
2. *Do not set the freezer temperatures lower than necessary (–18 °C is normal).*
3. *The ventilation grills should not be obstructed. This could decrease the freezing capability of the appliance.*
4. *Do not freeze lettuce, radishes, grapes, whole apples and pears, or fatty meat.*
5. *Do not salt or season fresh food before freezing.*
6. *Do not open the freezer door too often.*

Ignoring which of these six warnings could have caused the delay in the warning light going out?

Circle "Yes" or "No" for each of the six warnings.

Warning	Could ignoring the warning have caused a delay in the warning light going out?
Warning 1	Yes / No
Warning 2	Yes / No
Warning 3	Yes / No
Warning 4	Yes / No
Warning 5	Yes / No
Warning 6	Yes / No

Unit: *Freezer*
Question: *Question 2*[1]
Problem type: *Trouble shooting*
Item type: *Multiple choice*
Level: *Level 2*
PISA scale score: *573*
Item code: *X423Q02*

1. The numbering of the question here with Q2 preceding Q1 is only a reference to the order in which the questions were placed in the PISA assessment relative to how they were developed for the Field Trial. After the Field Trial, the order of the questions in the unit was reversed, but their original numbering was maintained for administrative purposes.

This question about the freezer problem called for students to diagnose the working of the warning light. Working through the three tests proposed, a student should indicate that moving the control to a warmer setting than the present one and getting the light to go off would potentially indicate that the freezer was still cooling down and had not yet reached the desired temperature at the present setting. Actions 1 and 3 would not provide information suggesting that the warning light was working properly. Getting these three answers correct is a Level 2 problem-solving task.

Response Coding guide for FREEZER Question 2

Full Credit

Code 1: No, Yes, No, in that order.

No Credit

Code 0: Other responses.

Code 9: Missing.

Unit: *Freezer* **Question:** *Question 1* **Problem type:** *Trouble shooting* **Item type:** *Multiple choice* **Level:** *Level 2* **PISA scale score:** *551* **Item code:** *X423Q01*

Another freezer question further extended the trouble shooting process, posing a series of six options with yes-no responses. Here the student is confronted with a series of warnings from the manual associated with freezer malfunctioning. The student is then asked to determine which of these warnings might be associated with a delay in the warning light turning off. This item was also judged to be at Level 2 as each of the decisions is based essentially on a single piece of information and its relationship to the mechanism. The question draws on a student's outside experience with freezers or similar appliances at a common sense level of knowledge that leads one to dismiss possible causes as irrelevant to a given situation.

Response coding guide for FREEZER question 1

Full Credit

Code 2: No, Yes, Yes, No, No, Yes, in that order.

Partial Credit

Code 1: One error.

No Credit

Code 0: Other responses.

Code 9: Missing.

Summary

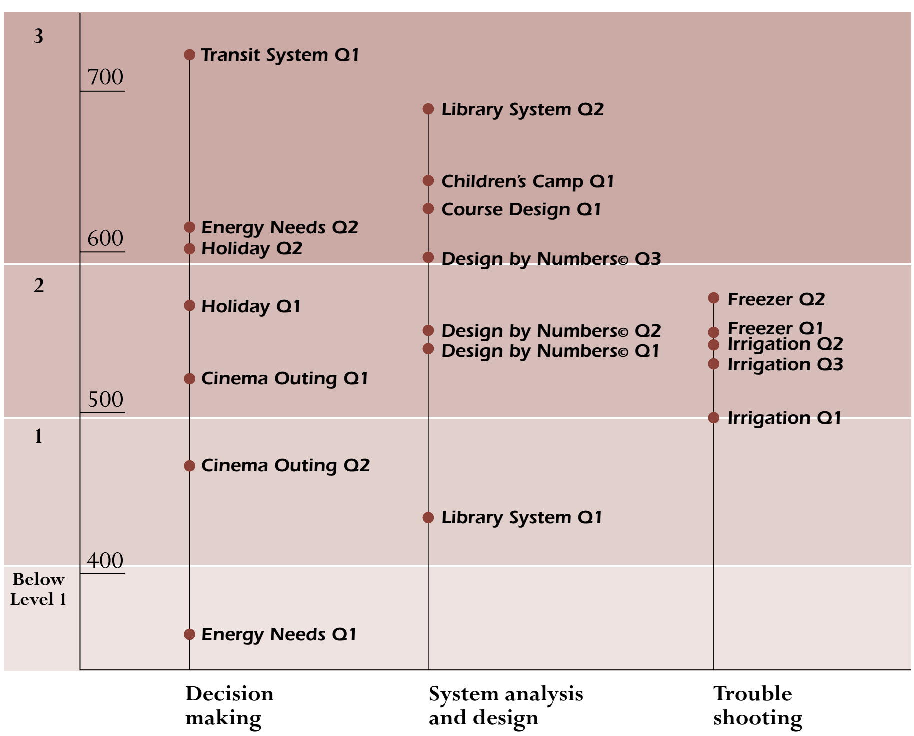
The 19 problems and 28 possible performance coding levels contained in the PISA cross-disciplinary problem-solving assessment provide a foundation for viewing students' problem solving in the three situations investigated: *decision making*, *system analysis and design*, and *trouble shooting*. Comparing the results with the model presented in Figure 2.1 of Chapter 2, shows how the characteristics of the problem types were carried through into the design of the stimulus materials and the items in each unit. Further, the potential sources of difficulty for items noted in the framework were indeed witnessed in the PISA problems in practice.

These items and the way students responded show how problems differ in difficulty and type.

Figure 4.13 shows the range of scores associated with full credit in the questions belonging to each of the three problem types. This shows a decreasing range of score values as one moves from *decision-making* to *system analysis and design* to *trouble-shooting* problems. There is also a convergence of item difficulty around the middle of the scale, rather than at its extremes, as one moves across the problem types.

Although some problem types have a wider difficulty range than others…

Figure 4.13 ■ **Graph of PISA problem-solving item scale values by problem type**



The analysis of the item performance indicates that the items between them cover the domain of problem solving as described in Chapter 2 and provide examples of student performance across the full range of the scale, from understanding problems to the solution of problems and the communication of results.

…overall they cover the domain as required.

The Role that Gender and Student Background Characteristics Play in Student Performance in Problem Solving

Introduction

This chapter looks at how both gender and student background relate to performance.

This chapter discusses how gender and student background characteristics relate to student performance in problem solving.

The PISA problem-solving tasks are intended to parallel situations in life and do not draw on specific curriculum knowledge. Therefore the effects of family, socio-economic and cultural background are particularly worth noting. Ideally, the future opportunities of any student should not depend on their socio-economic background. If students with less-advantaged backgrounds are less proficient at solving problems, they may risk difficulties in the transition to work or further education. These difficulties can then perpetuate social disparities from one generation to the next.

This chapter explores the relationships between problem-solving performance and a variety of student, family, and social factors. It compares gender differences in problem-solving performance and in other PISA assessment areas and then considers the impact of students' background characteristics on their problem-solving performance. These analyses include the occupational status of parents, the education of parents, "cultural" features and the immigration status of students and their parents. Most of these background variables have been shown to affect student performance in various assessment areas.

Gender differences in problem solving

Some recent studies have shown that females' historical disadvantage in mathematics performance is changing…

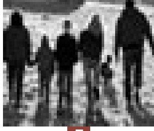
Given the importance education has on future opportunities in the life of an individual, all countries try to minimise gender-specific disadvantages for females or males in their education systems. Historically, this concern focused on gender-specific disadvantages affecting females. In recent studies, however, females have closed some gaps and even outperformed their male peers in some assessment areas.

…although males are still ahead in mathematics, but behind in reading.

Consequently, the underachievement of males has now also become a focus of educational research and policy development. Performance differences between female and male students found in recent international comparisons of student performance vary according to the assessment area. For instance, females generally outperformed males in reading while males tended to outperform females in mathematics (see also *Learning for Tomorrow's World – First Results from PISA 2003,* [OECD, 2004a]).

Are male strengths in mathematics reflected in a better general ability to solve problems?

It is not clear whether one should expect there to be a gender difference in problem solving. On the one hand, the questions posed in the PISA problem-solving assessment were not grounded in content knowledge, so males' or females' advantage in having mastered a particular subject area should not have shown through. On the other hand, as demonstrated in Chapter 3, there is a strong link between the analytical reasoning skills needed in mathematics and those needed in problem solving, and there is also a strong correlation between the results of these two PISA assessment areas. The extent to which the advantage

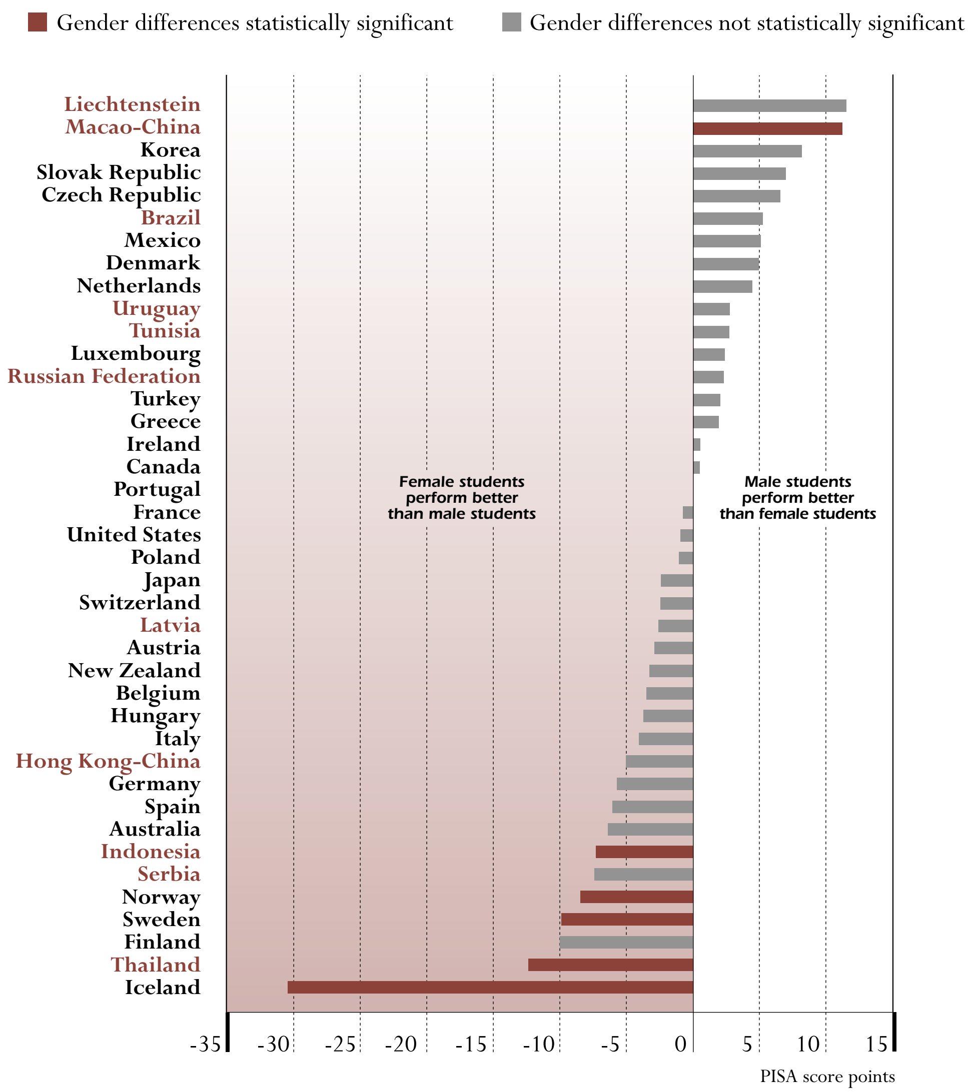


of male students in mathematics performance is replicated in problem solving may therefore give clues as to whether males do better in mathematics because they have mastered the subject better or because they have particular generic skills that help them solve mathematical problems.

Figure 5.1 shows the observed differences between the mean performance of female students and that of male students on the PISA problem-solving assessment. The length of the bars indicates the difference between genders on the problem-solving scale (to the right they show male students performing better, while to the left, they show female students performing better).

Figure 5.1 shows the differences between the mean performances of female and male students…

Figure 5.1 ■ **Gender differences in student performance in problem solving**

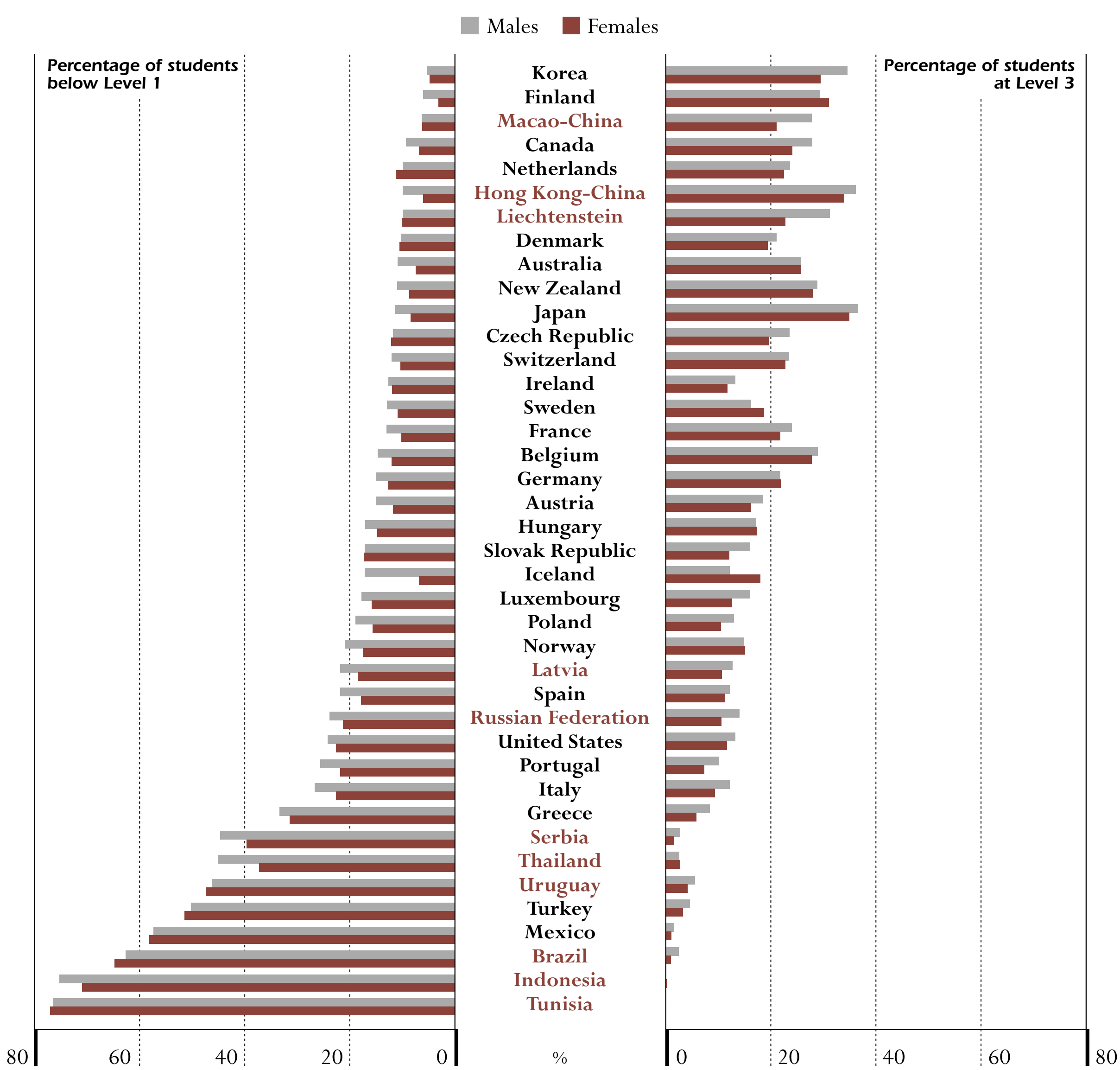


Countries are ranked in descending order of performance advantage for male students.
Source: OECD PISA 2003 database, Table 5.1.

...and indicates only minor gender differences in problem solving, with these slightly in favour of females overall.

Strikingly, few countries show statistically significant gender differences in problem solving. In Iceland, Norway and Sweden, as well as in the partner countries Indonesia and Thailand, female students outperform male students in problem solving. The partner country Macao-China is the only country where male students outperform female students in problem solving. As in mathematics and reading (see Table 5.1), in Iceland the advantage that female students have in problem solving is by far the largest compared to the other participating countries: female students score 30 points more than male students, representing a third of a proficiency level. However, in the remaining countries the largest gap in either direction is 12 score points or less.

Figure 5.2 ■ **Percentage of males and females performing below Level 1 and at Level 3 in problem solving**



Countries are ranked in ascending order of the percentage of males performing below Level 1 on the problem-solving scale.
Source: OECD PISA 2003 database, Table 5.1.

A review of gender differences at the highest and lowest proficiency levels…

Figure 5.2 shows the percentage of male and female students who are below Level 1 (*basic problem solvers*) and who are proficient at Level 3 (*reflective, communicative problem solvers*).

…shows that there are slightly more males among both the strongest and the weakest problem solvers.

This comparison shows that typically in participating countries, there are slightly more male students at the lowest and at the highest proficiency level. On average in OECD countries, 18 per cent of male students and 16 per cent of female students are below Level 1, while 19 per cent of male students and 18 per cent of female students reach Level 3. Thus, while male and female students do not differ markedly in their average problem-solving performance, the performance of male students overall is more spread out towards the extremes. This is more obviously so in some countries than in others – for example in Italy, a third more male than female students (12 per cent rather than 9 per cent) have the highest problem-solving skills, while a sixth more males (27 per cent rather than 23 per cent) lack basic problem-solving skills. Male students are not over-represented at both extremes in every country, but in all countries there are more male students in at least one of these categories. Moreover, a systematic measure of the variability in performance, the standard deviation, shows that male performance is more dispersed in every country except in the partner country Indonesia. On average in the OECD countries, the standard deviation for male students is 6 score points higher than for female students. This difference is most pronounced in Italy (the standard deviation for female students is 94 score points and the standard deviation for male students is 110 score points), as well as in the partner country Hong Kong-China (the standard deviation for female students is 90 score points and the standard deviation for male students is 104 score points). The standard deviation for male students is at least 12 score points higher than for female students in Poland, Portugal, Turkey and the partner country Liechtenstein.

Comparison with gender differences in other assessment areas

Male students' advantage in mathematics does not translate into stronger problem-solving skills…

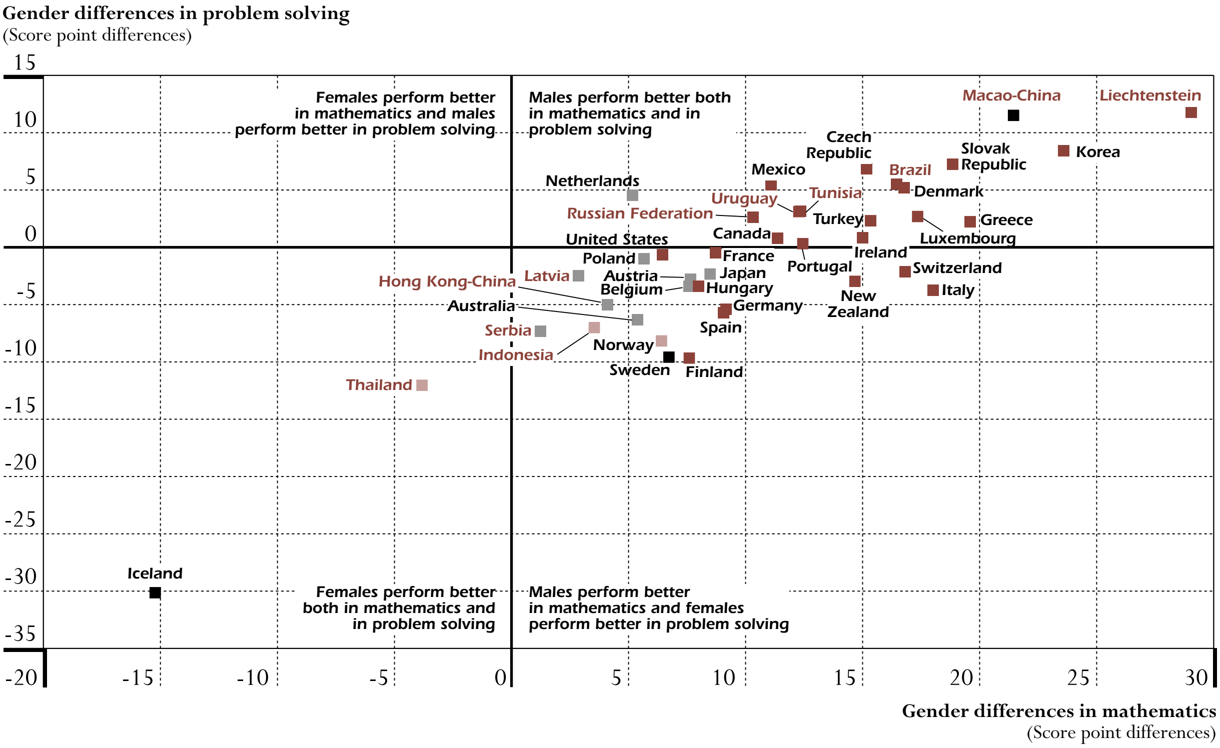
As reported in Chapter 3, performance in problem solving is closely related to performance in mathematics. The comparison of gender differences in these two assessment areas becomes particularly interesting since male students outperform female students in mathematics in most participating countries (see Table 5.1), whereas there are no pronounced gender differences in problem-solving performance. Figure 5.3 shows the relationship between gender differences in mathematics (horizontal axis) and gender differences in problem solving (vertical axis) among countries.

…although in many countries males do relatively better in the two assessment areas.

In half the countries gender differences are consistent in mathematics and in problem solving (see Figure 5.3). In countries with the largest advantage of male students in mathematics, such as Korea and the partner country Liechtenstein, male students also perform better than female students in problem solving (though not statistically significantly so). In Iceland and the partner country Thailand, where female students outperform male students in mathematics, as well as in countries with low gender differences in mathematics (*e.g.* the partner countries Indonesia, Latvia and Serbia[1]), female students outperform male students in problem solving (statistically significantly so in Iceland, Indonesia and Thailand).

Figure 5.3 ■ **Gender differences in problem solving and in mathematics**

- ■ Statistically significant gender differences in problem solving and mathematics
- ■ Statistically significant gender differences only in problem solving
- ■ Statistically significant gender differences only in mathematics
- ■ Gender differences in problem solving and mathematics are not statistically significant



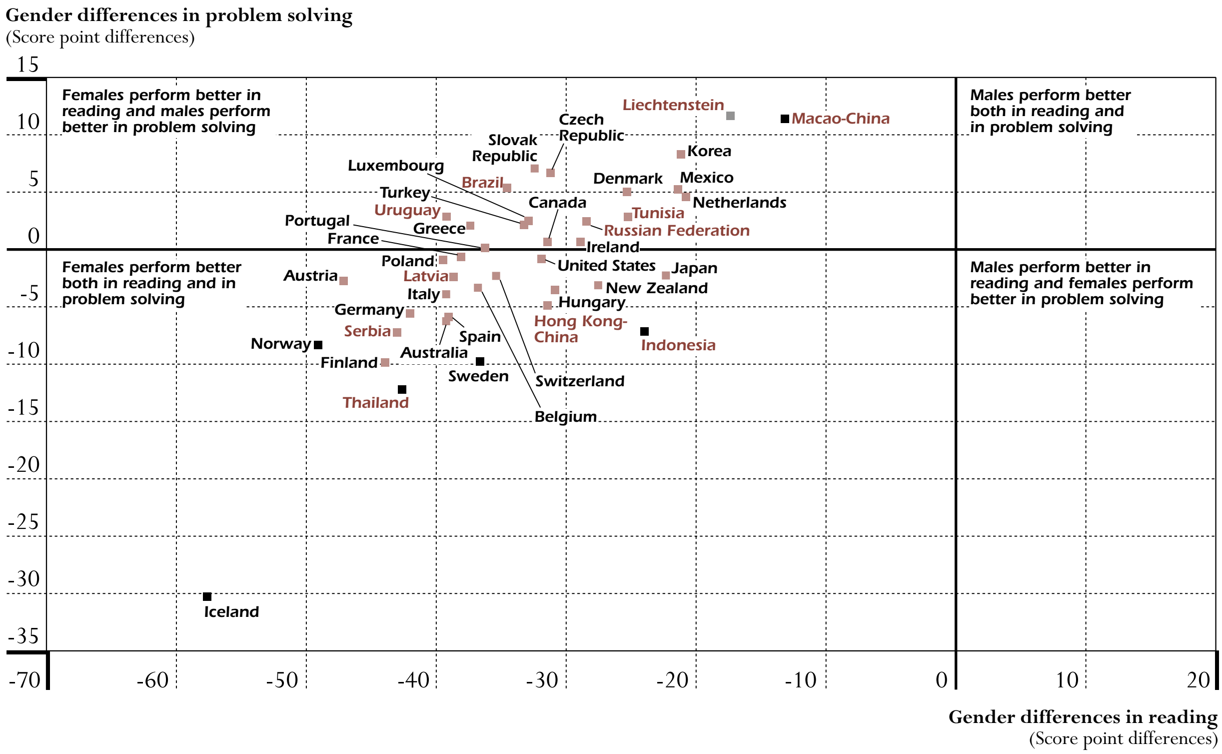
Positive values indicate that males perform better than females.
Source: OECD PISA 2003 database, Table 5.1.

Where female advantages in reading are strongest, female students also do relatively better in problem solving.

Figure 5.4 shows the relationship between gender differences in problem solving and gender differences in reading across countries. As when comparing problem solving to mathematics, the countries in which female students or male students do relatively better are similar across the two assessment areas. Even though the female reading advantage is not replicated in most countries for problem solving, where the female advantage in reading is relatively large such as Finland, Iceland and Norway, female students are also better problem solvers (statistically significantly so in Iceland and Norway). In Korea, the Netherlands and Mexico, as well as in the partner countries Liechtenstein and Macao-China, the advantage for female students in reading was relatively lower than in other countries and the advantage for female students disappears in problem solving (although the male advantage is only statistically significant in Macao-China). However, in Austria and Germany female students have a large advantage in reading, but their advantage in problem solving is small and not statistically significant.

Figure 5.4 ■ **Gender differences in problem solving and in reading**

■ Statistically significant gender differences in problem solving and reading
■ Statistically significant gender differences only in reading
■ Gender differences in problem solving and reading are not statistically significant



Positive values indicate that males perform better than females.
Source: OECD PISA 2003 database, Table 5.1.

Overall, a comparison of the different assessment areas shows gender differences to vary by area: Female students outperform their male peers in reading performance and male students reach somewhat higher levels of performance in mathematics. In contrast, in the majority of participating countries male and female students do not differ significantly in problem-solving performance, which is conceptualised to make cross-disciplinary demands on students' competencies. This may indicate that female and male students can draw on their own specific strengths when it comes to cross-disciplinary tasks. Male students' strengths in mathematics do not appear to derive from a superiority in analytical reasoning skills that has a disproportionate effect on general problem-solving abilities. Rather, gender-specific strengths seem to balance out in a way that leads to relatively equal outcomes for both genders in problem-solving performance. Moreover, the result may be viewed as an indication that in many countries there are no strong overall disadvantages for male students or female students as learners, but merely gender-specific strengths or preferences for certain subjects.

Thus it appears that neither gender is disadvantaged overall in problem solving, but each draws on their strengths.

Nevertheless, variations in gender differences remain important…

However, this does not mean that gender differences do not matter. First, it is important to note that, as in PISA 2000, the size of gender differences in reading in favour of female students is markedly larger than the gender differences in mathematics in favour of male students. In addition, the high consistency of the relative size of a country's gender differences across the PISA assessment areas indicates that in some countries there are still general advantages for one of the genders, *e.g.* for female students in Iceland.

…with some countries better at containing them than others.

Second, note that some countries have been much more successful than others in achieving equal performance for male and female students in each assessment area. In the Netherlands, for instance, the only significant performance difference is a comparatively small one of 21 score points in reading performance in favour of female students, while no significant differences are found in mathematics or problem solving. In Greece and Italy, on the other hand, large advantages for female students in reading (37 and 39 score points respectively) and large advantages for male students in mathematics (19 and 18 score points respectively) are found within the same countries. However, these subject-specific differences disappear in problem-solving performance: there is no significant gender difference for problem solving in Greece or Italy. PISA cannot show what is behind these performance differences, but it may be of interest to some education systems to encourage male and female students in the areas where they are significantly outperformed by the other gender.

Parental occupational status

Comparing students' problem-solving performance with parents' occupational status gives an important indicator of social disadvantage.

Parental occupational status, which is often closely interrelated with other attributes of socio-economic status, has a strong association with student performance. Generally, countries are interested in minimising such disparities. PISA problem-solving performance reflects the capability to deal with cross-disciplinary tasks that approximate real-life situations. Therefore it is a useful indicator of students' chances to successfully manage future challenges in life.

How much difference does parental occupational status make on average to student scores?

The average performance gap in problem solving between students in the top quarter of PISA's international socio-economic index of occupational status (whose parents have occupations in fields such as medicine, university teaching and law) and those in the bottom quarter (with occupations such as small-scale farming, truck-driving and serving in restaurants), amounts to an average of 76 score points, or four-fifths of a proficiency level in problem solving.[2] Expressed differently, one standard deviation (*i.e.* 16.4 units) on the PISA index of occupational status is associated with an average performance difference of 33 score points. Figure 5.5 shows the mean problem-solving performance for students in each quarter of the PISA index of occupational status. The length of the different lines represents the gap between students in the highest and lowest quarters of parental occupational status within each country.

Within OECD countries, students in the top national quarters on the international socio-economic index of occupational status reach a mean score of 542 score

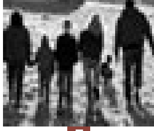
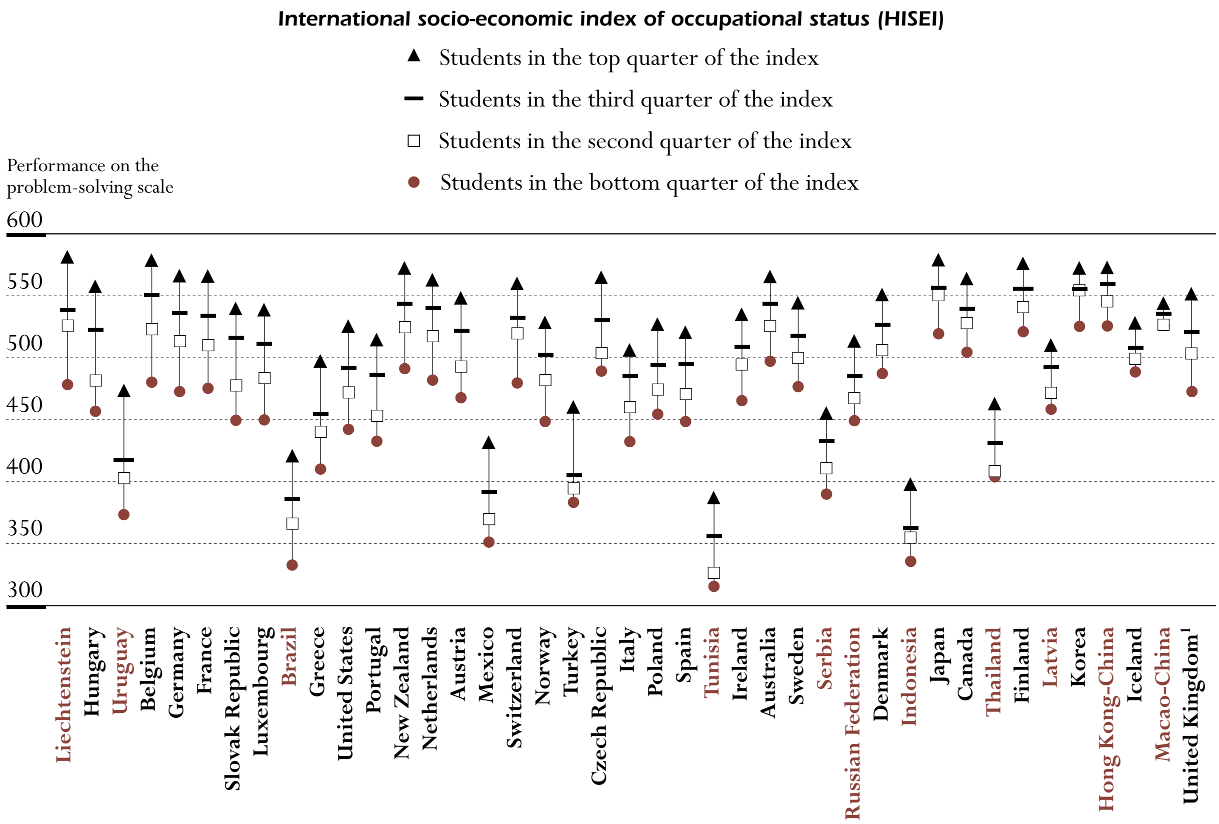


Figure 5.5 ■ **Parental occupational status and student performance in problem solving**



Countries are ranked in descending order of the difference in performance between students in the top and bottom quarters of the international socio-economic index of occupational status (HISEI).
1. Response rate too low to ensure comparability (see Annex A3).
Source: OECD PISA 2003 database, Table 5.2.

points on the problem-solving scale, or 42 score points above the OECD average. The average mean score in OECD countries for students in the bottom national quarters is only 465 score points. This means that students with parents in lower status occupations perform on average at the level of *basic problem solvers* (Level 1), while students with parents in higher status occupations perform on average at the level of *reasoning, decision-making problem solvers* (Level 2). As in previous OECD studies, this disadvantage associated with a low occupational status of students' parents is much more pronounced in some countries than in others. For instance, the difference between the problem-solving performance of the national top and bottom quarters on the index of parental occupation is equivalent to at least one proficiency level (94 score points) in Belgium (99 score points), Germany (94 score points) and Hungary (101 score points), as well as in the partner countries Liechtenstein (103 score points) and Uruguay (101 score points). In other countries this gap is limited to only half or less than half of a proficiency level (*e.g.* 40 score points in Iceland and 47 score points in Korea, as well as 47 score points in Hong Kong-China and 18 score points in Macao-China).

Figure 5.5 shows that it is equivalent to over one proficiency level in some countries, but less than half a proficiency level in others…

Furthermore, the analysis estimates the percentage of variability in student performance that can be predicted by the job that students' parents have. An amount of explained variance equal to zero means that there is no relationship between parental occupational status and problem-solving performance; an amount of 50 per cent of explained variance means that if one were to predict students' scores according to how well students with parents in similar occupations tend to perform, the result would show half of the variation in performance that is actually observed.

...and that the strength of the effect is much higher in some countries than in others.

Within OECD countries on average, 11 per cent of the variation in student performance is explained by parental occupational status (Table 5.2). This effect is significant in all of the participating countries and strongest in Belgium, Germany, Hungary, Portugal and the Slovak Republic, as well as in the partner countries Brazil, Liechtenstein and Uruguay (effects of between 34 and 41 score points; between 13 and 17 per cent explained variance in all countries except Uruguay [12 per cent]). In Iceland, Japan and Korea, and in the partner countries Hong Kong-China, Latvia and Macao-China, parents' occupation explains only 1 to 5 per cent of the variation in problem-solving performance (effects of between 12 and 24 score points).

Parental education

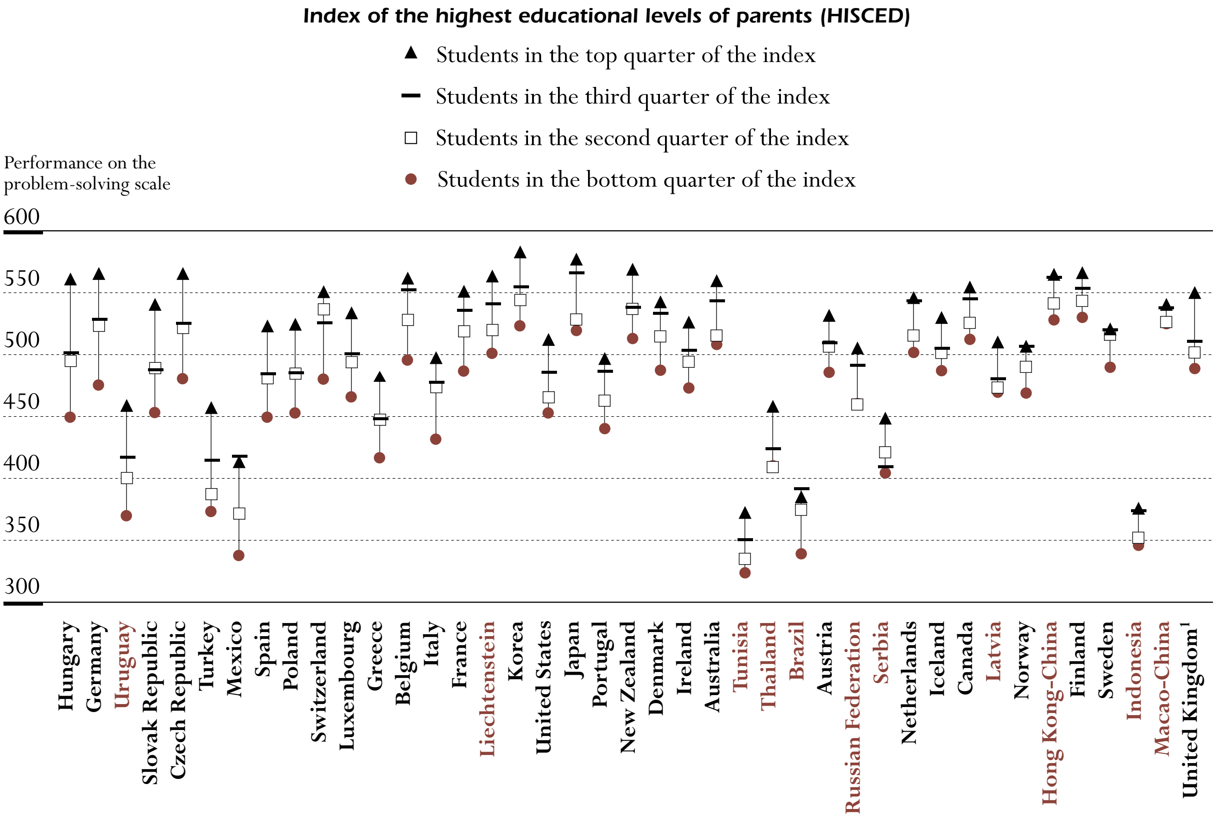
The quarter of students whose parents have the highest and lowest levels of education can be compared...

A strong predictor of student performance consistently found in previous studies is parental level of education. Parental level of education is classified in accordance with the International Standard Classification of Education, (ISCED, OECD, 1999). The higher level of the two parents is used as a single index for each student's parental level of education. The index for parental education is standardised for the OECD countries to have a mean of 0 and a standard deviation of 1 to facilitate interpretation of the results. Figure 5.6 illustrates the effect of parental education on problem-solving performance by displaying the mean performance of students in four groups ranked by the national values of this index. The length of the lines indicates the gap between students in the group whose parents have the highest and lowest levels of education in each country.

...and parental education can be seen to be a significant predictor across countries.

The level of parental education is a significant predictor of student performance in problem solving across all participating countries except in the partner country Macao-China. The effect of one standard deviation difference in parental education ranges from 11 score points in Portugal (the lowest for an OECD country) to 33 score points in Hungary (the OECD average is 20 score points). Parental education explains between 1 and 19 per cent of variance in problem-solving performance in all participating countries (the OECD average is 10 per cent). The effect is strongest in the Czech Republic, Hungary, Poland and the Slovak Republic (between 26 and 33 score points), and least pronounced in Finland and Portugal and the partner countries Brazil, Indonesia, Hong Kong-China and Tunisia (between 7 and 11 score points). Although the effect of parental education on problem-solving performance is rather low in

Figure 5.6 ■ **Parental education and student performance in problem solving**



Countries are ranked in descending order of the difference in performance between students in the top and bottom quarters of the index of highest educational level of parents (HISCED).

1. Response rate too low to ensure comparability (see Annex A3).

Source: OECD PISA 2003 database, Table 5.3.

the countries with highest mean performances (*e.g.* Finland and Korea, and the partner country Hong Kong-China), there is no consistent relationship between mean performance within each country and the effect of parental education across all participating countries.

Possessions related to "classical" culture

A similar analysis shows how many "classical" cultural possessions students have at home…

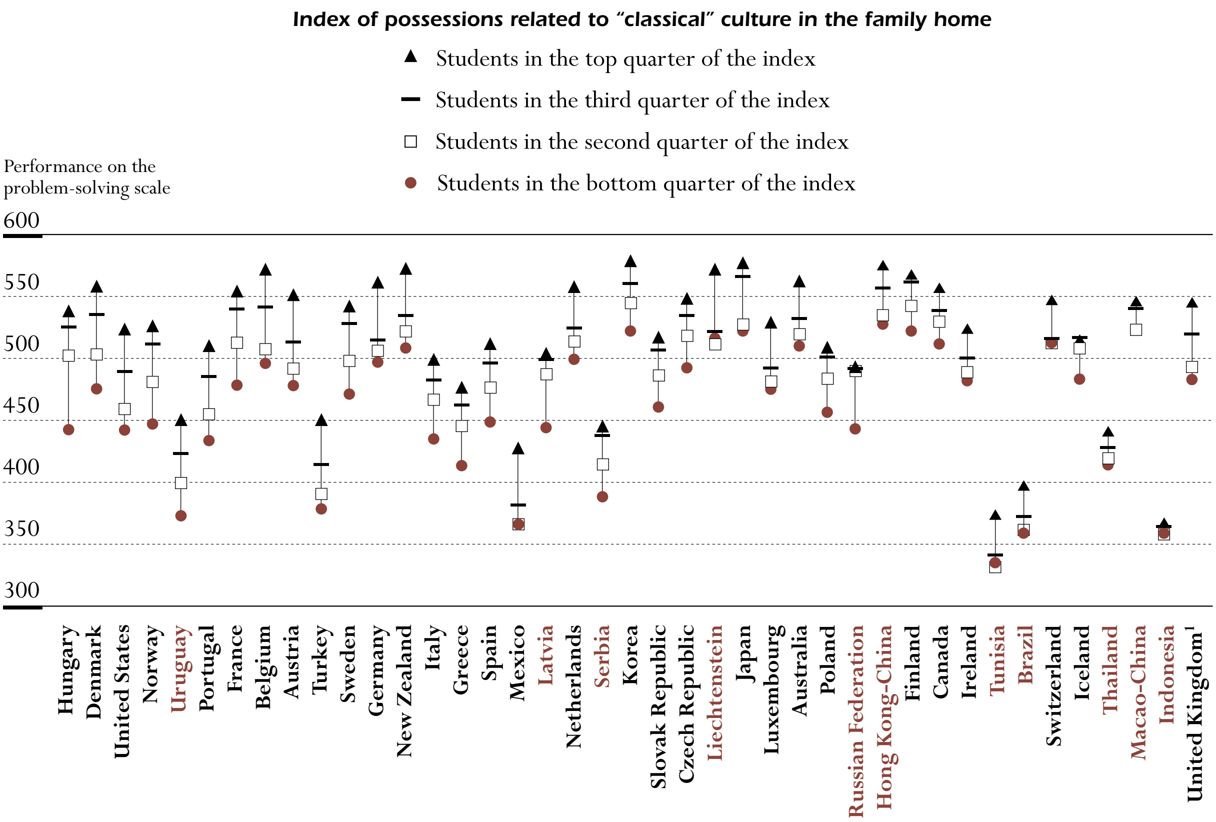
Like the level of parental education, the available possessions related to "classical" culture in a student's family is another background variable that was shown to be positively related to student performance in previous studies. To obtain an index of possessions in the family home related to "classical" culture, students in PISA 2003 were asked whether they had classical literature, books of poetry, and works of art in their homes. This index used is the same as in PISA 2000. As in PISA 2000, the highest levels of cultural possessions are found in Iceland, and the partner countries Latvia and the Russian Federation (see Chapter 4 in *Learning for Tomorrow's World – First Results from PISA 2003,* [OECD, 2004a]).

...and looks at the average scores of the quarters with the most and fewest of these possessions.

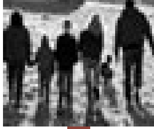
As for parental education, the index of cultural possessions was standardised for the OECD countries to have a mean of 0 and a standard deviation of 1. To illustrate the effects of cultural possessions on problem-solving performance, the population within each country is again divided into quarters ranked by the national values of this index. Figure 5.7 shows the mean problem-solving performance for students in the top and bottom of these groups. Thus, the length of the different lines demonstrates the gap between students with the highest and lowest levels of cultural possessions in each country.

Cultural possessions are significantly positively related to problem-solving performance in all participating countries. On average in the OECD countries, the performance gap for students with one unit more on the index of cultural possessions is 25 score points for problem-solving performance and accounts for 6 per cent of variation in student performance. This effect is strongest in Belgium, Denmark, Hungary, Mexico and the United States (effects of between 31 and 42 score points; between 9 and 17 per cent of variance explained)

Figure 5.7 ■ **Cultural possessions and student performance in problem solving**



Countries are ranked in descending order of difference between the top and bottom quarters on the index of possessions related to "classical" culture in the family home and performance on the problem-solving scale.
1. Response rate too low to ensure comparability (see Annex A3).
Source: OECD PISA 2003 database, Table 5.4.



and weakest in Canada and Switzerland, as well as in the partner countries Indonesia, Macao-China and Thailand (effects of between 5 and 18 score points; between 0.4 and 4 per cent of variance explained).

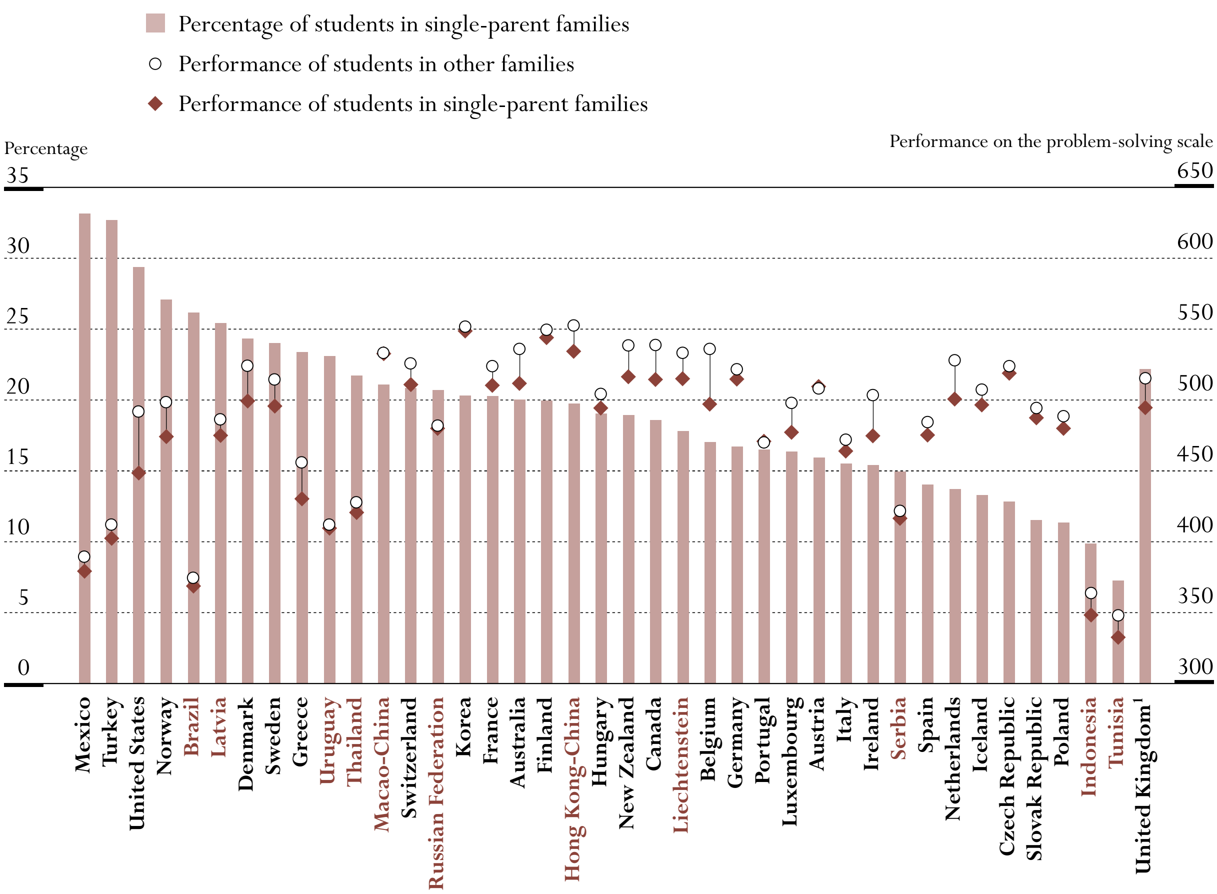
Family structure

Living with only one parent is often associated with lower performance…

The structure of a student's family is another background variable that can affect student performance. For example, students living with only one parent may receive less parental support with their learning compared to peers living with two parents. In PISA 2000, most distinct disadvantages in reading performance for students raised by a single parent were found in countries with high proportions of students in single-parent families. Figure 5.8 shows the percentage of students living in single-parent families and the mean problem-solving performance of students living in single-parent families and other family types.

Figure 5.8 ■ **Type of family structure and student performance in problem solving**

Percentage of students living in single-parent families (left scale) and performance of students on the problem-solving scale, by type of family (right scale)



Note: Countries are ranked in descending order of percentage of students in single-parent families.
1. Response rate too low to ensure comparability (see Annex A3).
Source: OECD PISA 2003 database, Table 5.5.

...although this disadvantage varies greatly and is not significant in half the PISA countries.

Within the OECD countries, between 11 and 33 per cent of students report to be living in a single-parent family (OECD average 19 per cent).[5] On average in the OECD countries, students living in such families score 23 points less than peers living with both parents or in another form of family with two guardians. Still, results show that this effect is not unavoidable. In 16 countries there is no significant disadvantage for students from single-parent families (*e.g.* Austria, Korea and Portugal). Figure 5.8 reveals no pronounced relationship between the proportion of students living in single-parent families and the disadvantage in problem-solving performance found for this group. The largest disadvantages of students living in single-parent families are found in the United States (44 score points) with one of the highest proportion of these students (29 per cent) and in Belgium (39 score points) with a relatively low proportion (17 per cent). In Mexico and Turkey, with the highest proportions of students reporting living in single-parent families, the disadvantage of these students in problem-solving performance is relatively small, although statistically significant (10 score points). A need to provide educational support to students from single-parent families can be seen in many countries, especially those with a high proportion of students living in single-parent families and a pronounced disadvantage of these students in problem-solving performance.

Place of birth and language spoken at home

Problem-solving performance can also be compared according to whether students and their parents are native to their country.

When families migrate from one country to another, their school-age children often find it hard to adjust to the new environment and often simultaneously are confronted with an unfamiliar language of instruction. In PISA, immigration status is assessed by asking students whether they and their parents were born in the country they are living in or in a different country. Based on these answers, students are classified into native students, who were born in and have parents who were born in the country of assessment, first-generation students, who were born in the country of assessment but whose parent(s) were born in a different country, and non-native students who were born in a different country. Figure 5.9 displays the percentages of first-generation and non-native students living within each country as well as the mean problem-solving performance of first-generation, non-native and native students.

First-generation and non-native students show on average a clear disadvantage, though its extent varies widely between countries.

Students classified as first-generation students and non-native students are on average at a clear disadvantage in terms of their problem-solving performance. On average in OECD countries, first-generation students score 26 points lower than native students and non-native students score 36 points lower. The differences in performance of students with an immigration background are not strongly related to the proportion of those among students within countries. In predominantly English-speaking countries with a relatively high proportion of students with an immigration background (Australia, Canada and the United States) and where 14 to 23 per cent of all 15-year-olds are non-native or first-generation students, the disadvantage of these groups is relatively small. In several European countries (Belgium, France, Germany and Switzerland) with between 12 and 20 per cent non-native or first-generation students, these groups

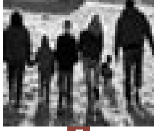
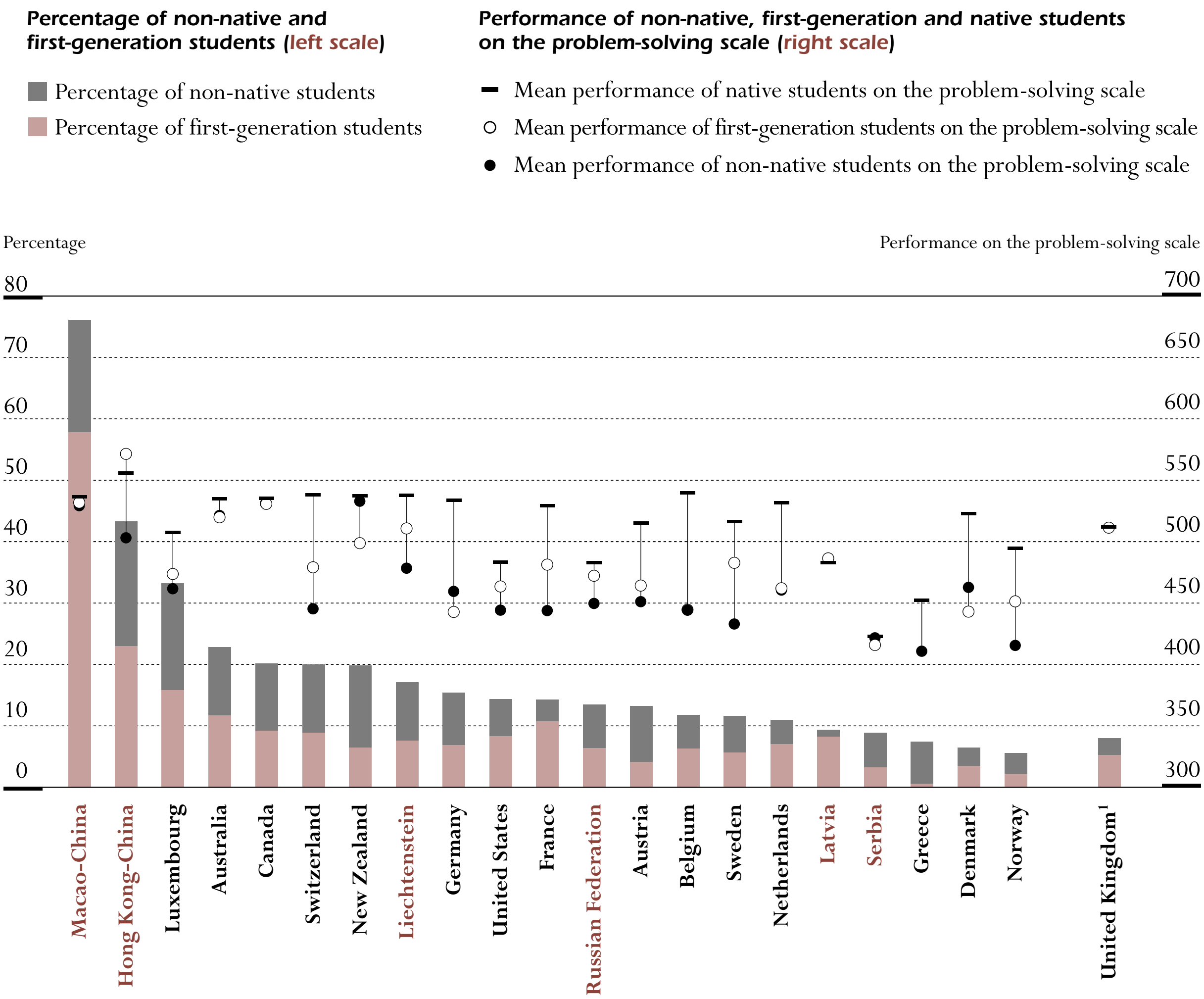


Figure 5.9 ■ **Place of birth and student performance in problem solving**



Note: Only countries with at least 3 per cent of students in at least one of these categories. Countries are ranked in descending order of the total percentage of non-native and first-generation students.
1. Response rate too low to ensure comparability (see Annex A3).
Source: OECD PISA 2003 database, Table 5.6.

perform distinctly less well than native students (a disadvantage of between 47 and 95 points). However, it should be noted that in some countries the mean performance of students with immigration backgrounds are based on very small numbers of students and should be interpreted carefully.

A similar comparison can be made according to the language spoken at home…

A consequence of immigration background may be that students speak a different language at home with their family than the language officially spoken in the country they live in and this is not limited to families with recent immigration history. Whatever the reason, students from families that speak a language at home that is different from the language of assessment or from other official languages or national dialects may experience difficulties in education due to a

relatively unfamiliar language of instruction. Figure 5.10 displays the percentage of students speaking a language at home that is different from the language of assessment or from other official languages or national dialects within countries as well as the mean problem-solving performance of these students and their peers who speak the same language at home.

...with those not speaking the official language more than one proficiency level behind in some countries.

On average in the OECD countries, students speaking a different language at home score 39 points lower in problem solving than their peers speaking the language of assessment or another official language or national dialect at home. Again, this disadvantage is not related to the percentage of students not speaking an official language within each country. While the disadvantage of this group is

Figure 5.10 ■ **Home language and student performance in problem solving**

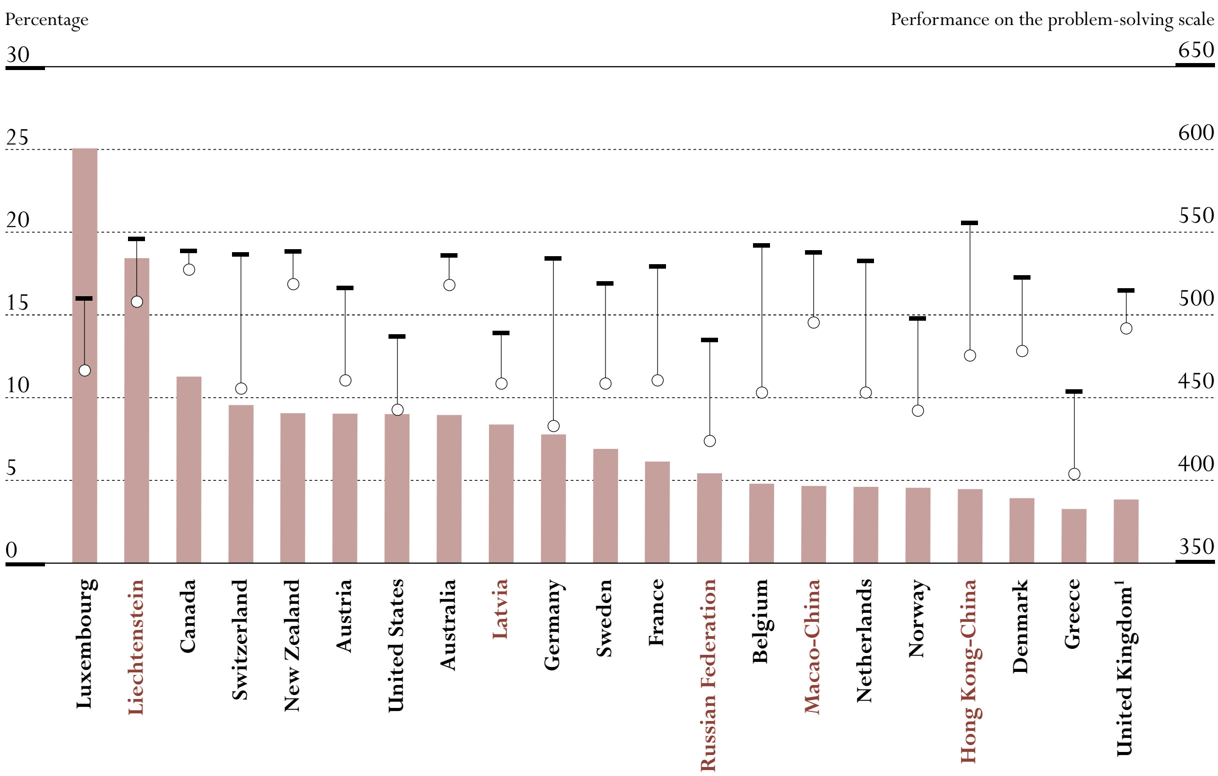
On the left scale:

■ Percentage of students who speak a language at home most of the time that is different from the language of assessment, from other official languages or from other national dialects

On the right scale:

▬ Mean performance on the problem-solving scale of students who speak a language at home most of the time that is the same as the language of assessment, other official languages or other national dialects

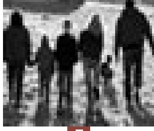
○ Mean performance on the problem-solving scale of students who speak a language at home most of the time that is different from the language of assessment, from other official languages or from other national dialects



Note: Only countries with at least 3 per cent of students in at least one of these categories. Countries are ranked in descending order of students who speak a language at home most of the time that is different from the language of assessment, from other official languages or from national dialects.

1. Response rate too low to ensure comparability (see Annex A3).

Source: OECD PISA 2003 database, Table 5.7.



relatively small in Australia and Canada (18 and 11 score points respectively), it is about one international standard deviation in Germany and Switzerland (101 and 81 score points respectively). In all of these countries between 8 and 11 per cent of 15-year-old students do not speak an official language at home.

This suggest that linguistic factors and migration interact to disadvantage some students.

Effects of the place of birth and the language spoken at home on problem-solving performance certainly reflect difficulties students with migration background experience within education. These disadvantages can perpetuate existing socio-economic gaps between the native population and immigrants across generations and should therefore be of concern for policy makers, especially in countries with a substantial number of immigrants. There are many different factors behind the relationship between students with a migration background and lower problem-solving performance, such as the previous educational experiences of these students, the unfamiliar language of instruction, the change from one education system to another, or other reasons. *Learning for Tomorrow's World – First Results from PISA 2003* (OECD, 2004a) examines some of these factors in more detail, and in particular, how they interact with other socio-economic characteristics.

Implications for policy

Prevailing issues in mathematics and science education over the past 50 years have been the eradication of gender differences in schooling and in re-examining cultural views about females' roles in these two subject areas. Vast improvements have been made in narrowing the gaps between male and female performance in school and in curriculum-related assessments in mathematics and science. However, little is directly known about gender differences in cross-curricular problem-solving performance when the emphasis is on real-world problems. The PISA 2003 assessment provides a window into the comparison of gender-related performance for 15-year-olds.

The fact that neither males nor females are systematically better at problem solving makes this domain a useful indicator of countries' gender bias.

In contrast to performances in mathematics and reading, there are no consistent differences in the problem-solving performances of male and female students. This may indicate that gender specific strengths or preferences for certain subjects can be compensated for when solving cross-disciplinary tasks. In this sense, problem solving provides also a good overall indicator of educational outcomes for males compared with those of females in an individual country, and hence of the extent to which societies have removed gender-based disadvantages in cognitive performance. More generally, one could regard problem solving as an area not affected by particular characteristics of one part of the curriculum that may favour one group over another, and thus for example as a more neutral indicator of the extent of differences in opportunity based on student social background.

Overall males are more likely to be found at the lower and upper ends of the performance distribution.

While male and female students do not differ markedly in their average performance, the variability of problem-solving performance is larger among males than females. More male students are found at the lower end of the performance distribution and consequently may face considerably restricted

opportunities in future life – a problem that warrants attention by policy makers. In addition, there may be disadvantages for females related to their under-representation at the highest performance level.

As in mathematics performance, students' family backgrounds have an important effect on their general problem-solving capacities …

Most effects of background variables on problem-solving performance presented in this chapter are very similar to effects of the examined variables on performances in other assessment areas. This shows that students from less advantaged backgrounds are disadvantaged not only in relation to how well they pick up the school curriculum, but also in terms of their acquisition of general problem-solving skills. Countries should be concerned that social background has such a strong effect not just on curricular outcomes but also on acquisition of general skills. Many studies are pointing to the importance of employees acquiring problem-solving skills in the modern workplace (*e.g.* US Department of Labour, 1991, McCurry, 2002, ILO, 1998, OECD 2001b). In particular, employees in modern firms need to be able to participate effectively in problem-solving groups in which a cross-section of employees work together to streamline and improve the workplace in terms of efficiency and productivity.

…and educational strategies need to address this.

The finding that inequities related to socio-economic and cultural background variables are not restricted to performance differences in school-related tasks underlines the importance of policy makers looking for strategies to raise problem-solving competence among disadvantaged groups.

Notes

1. For the country Serbia and Montenegro, data for Montenegro are not available. The latter accounts for 7.9 per cent of the national population. The name "Serbia" is used as a shorthand for the Serbian part of Serbia and Montenegro.

2. Father's or mother's occupation was used for this comparison, whichever was higher on the PISA international socio-economic index of occupational status (HISEI).

3. Japan is excluded from the following analyses because of a high proportion of missing data.

REFERENCES

Carroll, J. B. (1996), "Mathematical abilities: Some results from factor analysis", R. J. Sternberg, & T. Ben Zeev (eds.), *The Nature of Mathematical Thinking,* Lawrence Erlbaum Associates, Mahwah, N.J., 3-25.

Ganzeboom, H.B.G., P.M. De Graaf and **D.J. Treiman** (1992), "A standard international socio-economic index of occupational status", Social Science Research, Vol. 21, Issue 1, Elsevier Ltd., 1-56.

Green, F., D. Ashton, B. Burchell, **B. Davies**, and **A. Felkstead** (1997), "An Analysis of Changing Work Skills in Britain", paper presented at the Analysis of Low Wage Employment Conference, Centre for Economic Performance, London School of Economics, 12-13 December.

International Labour Office (1998), *World Employment Report 1998-99 Employability in the Global Economy – How Training Matters,* International Labour Office, Geneva.

Johnson, A. H. (2000), *Changing Skills for a Changing World: Recommendations for Adult Literacy Policy In Aotearoa/New Zealand*, Department of Labour Occasional Paper 2000, no. 2, Wellington, New Zealand.

Lerman, R. I., and **F. Skidmore** (1999), *Helping Low-Wage Workers: Policies for the Future,* Urban Institute, Washington, D.C.

McCurry, D. (2002), *Notes towards an overarching model of cognitive abilities,* unpublished report, Australian Council for Educational Research, Camberwell, Victoria.

OECD (1999), *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries,* OECD, Paris.

OECD (2001a), *Knowledge and Skills for Life – First Results from PISA 2000,* OECD, Paris.

OECD (2002a), *Reading for Change – Performance and Engagement across Countries,* OECD, Paris.

OECD (2002b), *PISA 2000 Technical Report,* OECD, Paris.

OECD (2003a), *Literacy Skills for the World of Tomorrow – Further Results from PISA 2003,* OECD, Paris.

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

OECD (2003c), *Learners for Life – Student Approaches to Learning,* OECD, Paris.

OECD (2003d), *Student Engagement at School – A Sense of Belonging and Participation,* OECD, Paris.

OECD (2003e), *Education at a Glance,* OECD, Paris.

OECD (2004a), *Learning for Tomorrow's World – First Results from PISA 2003,* OECD, Paris.

OECD (2004b), *What Makes School Systems Perform,* OECD, Paris.

OECD (forthcoming), *School Factors Relating to Quality and Equity,* OECD, Paris.

OECD (forthcoming), *PISA 2003 Technical Report,* OECD, Paris.

Steedman, H. (1999) "Low Skills: How the Supply is Changing across Europe", in *European Trends in Occupations and Qualifications,* CEDEFOP, Luxembourg.

United States Department of Labour (1991), The Secretary's Commission on Achieving Necessary Skills (SCANS), *What Work Requires of Schools: A SCANS Report for America 2000,* United States Department of Labor, Washington, D.C.

Warm, T.A. (1985), "Weighted maximum likelihood estimation of ability in Item Response Theory with tests of finite length", *Technical Report CGI-TR-85-08*, U.S. Coast Guard Institute, Oklahoma City.

Workbase: The New Zealand Centre for Workforce Literacy Development (2000), *Statement of Objectives,* Workbase: The New Zealand Centre for Workforce Literacy Development, Auckland.

Annex A

TECHNICAL BACKGROUND

Annex A1: Construction of indices and other derived measures from the student context questionnaire

This section explains the indices derived from the student and school context questionnaires that are used in this report.

Several of PISA's measures reflect indices that summarise responses from students or school representatives (typically principals) to a series of related questions. The questions were selected from larger constructs on the basis of theoretical considerations and previous research. Structural equation modelling was used to confirm the theoretically expected behaviour of the indices and to validate their comparability across countries. For this purpose, a model was estimated separately for each country and collectively for all OECD countries.

For a detailed description of other PISA indices and details on the methods see the *PISA 2000 Technical Report* (OECD, 2002d) or the *PISA 2003 Technical Report* (OECD, forthcoming).

Unless otherwise indicated, where an index involves multiple questions and student responses, the index was scaled using a weighted maximum likelihood estimate (WLE) (see Warm, 1985), using a one-parameter item response model, which in the case of items with more than two categories was the Partial Credit Model. The scaling was done in three stages:

- The item parameters were estimated from equal-sized sub-samples of students from each OECD country.
- The estimates were computed for all students and all schools by anchoring the item parameters obtained in the preceding step.
- The indices were then standardised so that the mean of the index value for the OECD student population was zero and the standard deviation was one (countries being given equal weight in the standardisation process).

To illustrate the meaning of the international scores on the index, item maps were constructed that relate the index value to typical student responses to the questions asked. These item maps can be found on the website *www.pisa.oecd.org*. The vertical lines on the maps indicate for each of the index scores at the top of the figure which response a student is most likely to give, with zero representing the average student response across OECD countries.

It is important to note that negative values for an index do not necessarily imply that students responded negatively to the underlying questions. A negative value merely indicates that a group of students (or all students, collectively, in a single country) or principals responded less positively than all students or principals did on average across OECD countries. Likewise, a positive value on an index indicates that a group of students or principals responded more favourably, or more positively, than students or principals did, on average, in OECD countries.

Terms enclosed in brackets < > in the following descriptions were replaced in the national versions of the student and school questionnaires by the appropriate national equivalent. For example, the term <qualification at ISCED level 5A> was translated in the United States into "Bachelor's degree, post-graduate certificate program, Master's degree program or first professional degree program". Similarly the term <classes in the language of assessment> in Luxembourg was translated into "German classes" or "French classes" depending on whether students received the German or French version of the assessment instruments.

For additional information on how these indices were constructed, see the *PISA 2000 Technical Report* (OECD, 2002b) or the *PISA 2003 Technical Report* (OECD, forthcoming).

Student level variables

Student background

Family structure

Students were asked to report who usually lived at home with them. The response categories were then grouped into four categories: *i)* ***single-parent family*** (students who reported living with one of the following: mother, father, female guardian or male guardian); *ii)* ***nuclear family*** (students who reported living with a mother and a father); *iii)* ***mixed family*** (students who reported living with a mother and a guardian, a father and a guardian, or two guardians); and *iv)* ***other response combinations***. Non-responses are maintained as missing.

Parental occupations

Students were asked to report their mothers' and fathers' occupations, and to state whether each parent was in full-time paid work; part-time paid work; not working but looking for a paid job; or "other".

The open-ended responses for occupations were then coded in accordance with the International Standard Classification of Occupations (ISCO 1988).

The PISA ***international socio-economic index of occupational status*** (ISEI) was derived from students' responses on parental occupation. The index captured the attributes of occupations that convert parents' education into income. The index was derived by the optimal scaling of occupation groups to maximise the indirect effect of education on income through occupation and to minimise the direct effect of education on income, net of occupation (both effects being net of age). For more information on the methodology, see Ganzeboom *et al.* (1992). The ***highest international socio-economic index of occupational status*** (HISEI) corresponds to the highest ISEI of either the father or the mother.

The variables on students' fathers' and mothers' occupations were also transformed into four ***socio-economic categories****:* *i)* white-collar high-skilled: legislators, senior officials and managers, professionals, technicians and associate professionals; *ii)* white-collar low-skilled: service workers, shop and market sales workers and clerks; *iii)* blue-collar high-skilled: skilled agricultural and fishery workers and craft and related trades workers; and *iv)* blue-collar low-skilled: plant and machine operators and assemblers and elementary occupations.

Educational level of parents

Parental education is a family background variable that is often used in the analysis of educational outcomes. Indices were constructed using information on the ***educational level of the father,*** the ***educational level of the mother,*** and the highest level of education between the two parents, referred to as the ***highest educational level of parents***. Students were asked to identify the highest level of education of their mother and father on the basis of national qualifications, which were then coded in accordance with the International Standard Classification of Education (ISCED 1997, see OECD, 1999b) in order to obtain internationally comparable categories of educational attainment. The resulting categories were: (0) for no education; (1) for the completion of <ISCED Level 1> (primary education); (2) for completion of <ISCED Level 2> (lower secondary education); (3) for the completion of <ISCED Level 3B or 3C> (vocational/pre-vocational upper secondary education, aimed in most countries at providing direct entry into the labour market); (4) for completion of <ISCED Level 3A> (upper secondary education, aimed in most countries at gaining entry into tertiary-type A [university level] education) and/or <ISCED Level 4> (non-tertiary post-secondary); (5) for qualifications in <ISCED 5B> (vocational tertiary); and (6) for completion of<ISCED Level 5A, 6> (tertiary-type A and advanced research programmes).

Immigration background

The index on ***immigrant background*** was derived from students' responses to questions about whether or not their mother and their father were born in the country of assessment or in another country. The response categories were then grouped into three categories: *i)* "native" students (those students born in the country of assessment or who had at least one parent born in that country); *ii)* "first-generation" students (those born in the country of assessment but whose parents were born in another country); and *iii)* "non-native" students (those born outside the country of assessment and whose parents were also born in another country). For some comparisons, first-generation and non-native students were grouped together.

Language used at home

Students were asked if the language spoken at home most of the time or always was the language of assessment, another official national language, other national dialect or language, or another language. The index on ***language spoken at home*** distinguishes between students who report using the language of assessment, another official national language, a national dialect or another national language always or most of the time at home and those who report using another language always or most of the time at home.

In most countries, the languages were individually identified and were coded internationally to allow for further research and analysis in this area.

Possessions related to "classical" culture in the family home

The PISA index of ***possessions related to "classical" culture in the family home*** was derived from students' reports on the availability of the following items in their home: classic literature (examples were given), books of poetry and works of art (examples were given). Scale construction was performed through IRT scaling and positive values indicate higher levels of cultural possessions.

Annex A2: Detailed results from the factor analysis in Chapter 3

Method used for the factor analysis

The extraction method for the factor analysis was principal components analysis, with the Oblimin rotation method to allow for the factors to be correlated.

Eigenvalues for the first 12 factors

The eigenvalues of the first factors are shown in Table A2.1.

Table A2.1
Eigenvalues of the first 12 factors and total variance explained

Component	Initial Eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings[1]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	9.768	7.631	7.631	9.768	7.631	7.631	8.998
2	3.689	2.882	10.513	3.689	2.882	10.513	6.420
3	3.668	2.866	13.379				
4	3.390	2.648	16.027				
5	3.259	2.546	18.573				
6	3.049	2.382	20.955				
7	3.029	2.367	23.322				
8	2.862	2.236	25.558				
9	2.714	2.120	27.678				
10	2.667	2.083	29.762				
11	2.607	2.037	31.798				
12	1.497	1.169	32.968				

Extraction Method: Principal Component Analysis.
1. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

It is clear that there is one dominant factor with an eigenvalue of 9.8, followed by a group of 10 factors with eigenvalues between 2.6 and 3.7. To simplify the interpretations of the factor structure, two factors were chosen for the rotated solution. The rationale for the selection of two factors was not based on statistical criteria. Rather, it was based on a hypothesis that mathematics items and reading items should load separately on the first two factors, and it was of interest to see how the problem-solving items loaded on these two dimensions.

Component correlation matrix

Table A2.2 shows the component correlation matrix between the two factors.

Table A2.2
Component correlation matrix

Component	1	2
1	1.000	.362
2	.362	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

A cautionary note about the exploratory factor analysis

The results of the factor analysis are based on the analysis of correlations between variables at the student level. Due to the clustered structure of the PISA sample, these correlations reflect not only relations on the individual level but also heterogeneity between different schools. This means that correlations in terms of relations at individual level may be inflated due to general performance differences between students in different schools, countries etc. These inflated correlations will in turn lead to a more homogeneous factor structure – *i.e.* fewer dimensions – than one that would probably emerge from correlations where the clustered sampling is taken into account.

Overall, this analysis was not aimed at finding distinct cognitive factors in the PISA test. Rather, it was undertaken to provide some indication of the relative relationships between problem-solving items and mathematics and reading items. To this end, the results of the analysis showed clear indications that problem-solving items were more closely related to the general factor, and to mathematics items, than to reading items.

Annex A3: The PISA target population and the PISA samples

The PISA concept of yield and the definition of the PISA target population

PISA 2003 provides an assessment of the cumulative yield of education and learning at a point at which most young adults are still enrolled in initial education.

A major challenge for an international survey is to operationalise such a concept in ways that guarantee the international comparability of national target populations.

Differences between countries in the nature and extent of pre-primary education and care, the age of entry to formal schooling and the institutional structure of educational systems do not allow the definition of internationally comparable grade levels of schooling. Consequently, international comparisons of educational performance typically define their populations with reference to a target age group. Some previous international assessments have defined their target population on the basis of the grade level that provides maximum coverage of a particular age cohort. A disadvantage of this approach is that slight variations in the age distribution of students across grade levels often lead to the selection of different target grades in different countries, or between education systems within countries, raising serious questions about the comparability of results across, and at times within, countries. In addition, because not all students of the desired age are usually represented in grade-based samples, there may be a more serious potential bias in the results if the unrepresented students are typically enrolled in the next higher grade in some countries and the next lower grade in others. This would exclude students with potentially higher levels of performance in the former countries and students with potentially lower levels of performance in the latter.

In order to address this problem, PISA uses an age-based definition for its target population, *i.e.* a definition that is not tied to the institutional structures of national education systems: PISA assesses students who were aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months at the beginning of the assessment period and who were enrolled in an educational institution, regardless of the grade levels or type of institution in which they were enrolled, and regardless of whether they were in full-time or part-time education (15-year-olds enrolled in Grade 6 or lower were excluded from PISA 2003, but, among the countries participating in PISA 2003, such students only exist in significant numbers in Brazil). Educational institutions are generally referred to as ***schools*** in this publication, although some educational institutions (in particular some types of vocational education establishments) may not be termed schools in certain countries. As expected from this definition, the average age of students across OECD countries was 15 years and 8 months, a value which varied by less than 0.2 years between participating countries.

As a result of this population definition, PISA makes statements about the knowledge and skills of a group of individuals who were born within a comparable reference period, but who may have undergone different educational experiences both within and outside schools. In PISA, these knowledge and skills are referred to as the ***yield*** of education at an age that is common across countries. Depending on countries' policies on school entry and promotion, these students may be distributed over a narrower or a wider range of grades. Furthermore, in some countries, students in PISA's target population are split between different education systems, tracks or streams.

If a country's scale scores in problem solving are significantly higher than those in another country, it cannot automatically be inferred that the schools or particular parts of the education system in the first country are more effective than those in the second. However, one can legitimately conclude that the cumulative impact of learning experiences in the first country, starting in early childhood and up to the age of 15 and embracing experiences both in school and at home, have resulted in higher outcomes in this PISA assessment area.

The PISA target population did not include residents attending schools in a foreign country.

To accommodate countries that desired grade-based results for the purpose of national analyses, PISA 2003 provided an international option to supplement age-based sampling with grade-based sampling.

Population coverage

All countries attempted to maximise the coverage of 15-year-olds enrolled in education in their national samples, including students enrolled in special educational institutions. As a result, PISA 2003 reached standards of population coverage that are unprecedented in international surveys of this kind.

Table A3.1
PISA target populations and samples

	Population and sample information				
	(1) Total population of 15-year-olds	(2) Total enrolled population of 15-year-olds at grade 7 or above	(3) Total in national desired target population	(4) Total school-level exclusions	(5) Total in national desired target population after all school exclusions and before within-school exclusions
OECD countries					
Australia	268 164	250 635	248 035	1 621	246 414
Austria	94 515	89 049	89 049	321	88 728
Belgium	120 802	118 185	118 185	561	117 624
Canada	398 865	399 265	397 520	6 600	390 920
Czech Republic	130 679	126 348	126 348	1 294	125 054
Denmark	59 156	58 188	58 188	628	57 560
Finland	61 107	61 107	61 107	1 324	59 783
France	809 053	808 276	774 711	18 056	756 655
Germany	951 800	916 869	916 869	5 600	911 269
Greece	111 286	108 314	108 314	808	107 506
Hungary	129 138	123 762	123 762	3 688	120 074
Iceland	4 168	4 112	4 112	26	4 086
Ireland	61 535	58 997	58 906	864	58 042
Italy	561 304	574 611	574 611	2 868	571 743
Japan	1 365 471	1 328 498	1 328 498	13 592	1 314 906
Korea	606 722	606 370	606 370	2 729	603 641
Luxembourg	4 204	4 204	4 204	0	4 204
Mexico	2 192 452	1 273 163	1 273 163	46 483	1 226 680
Netherlands	194 216	194 216	194 216	2 559	191 657
New Zealand	55 440	53 293	53 160	194	52 966
Norway	56 060	55 648	55 531	294	55 237
Poland	589 506	569 294	569 294	14 600	554 694
Portugal	109 149	99 216	99 216	826	98 390
Slovak Republic	84 242	81 945	81 890	1 042	80 848
Spain	454 064	418 005	418 005	1 639	416 366
Sweden	109 482	112 258	112 258	1 615	110 643
Switzerland	83 247	81 020	81 020	2 760	78 260
Turkey	1 351 492	725 030	725 030	5 328	719 702
United Kingdom	768 180	736 785	736 785	24 773	712 012
United States	3 979 116	3 979 116	3 979 116	0	3 979 116
Partner countries					
Brazil	3 618 332	2 359 854	2 348 405	0	2 348 405
Hong Kong-China	75 000	72 631	72 631	601	72 030
Indonesia	4 281 895	3 113 548	2 968 756	9 292	2 959 464
Latvia	37 544	37 138	37 138	1 419	35 719
Liechtenstein	402	348	348	0	348
Macao-China	8 318	6 939	6 939	0	6 939
Russian Federation	2 496 216	2 366 285	2 366 285	23 445	2 342 840
Serbia	98 729	92 617	92 617	4 931	87 686
Thailand	927 070	778 267	778 267	7 597	770 670
Tunisia	164 758	164 758	164 758	553	164 205
Uruguay	53 948	40 023	40 023	59	39 964

	Population and sample information				
	(6) Percentage of all school-level exclusions	(7) Number of participating students	(8) Weighted number of participating students	(9) Number of excluded students	(10) Weighted number of excluded students
OECD countries					
Australia	0.65	12 551	235 591	228	3 612
Austria	0.36	4 597	85 931	60	1 099
Belgium	0.47	8 796	111 831	102	1 193
Canada	1.66	27 953	330 436	1 993	18 328
Czech Republic	1.02	6 320	121 183	22	218
Denmark	1.08	4 218	51 741	214	2 321
Finland	2.17	5 796	57 883	79	725
France	2.33	4 300	734 579	51	8 158
Germany	0.61	4 660	884 358	61	11 533
Greece	0.75	4 627	105 131	144	2 652
Hungary	2.98	4 765	107 044	62	1 065
Iceland	0.63	3 350	3 928	79	79
Ireland	1.47	3 880	54 850	139	1 619
Italy	0.50	11 639	481 521	188	6 794
Japan	1.02	4 707	1 240 054	0	0
Korea	0.45	5 444	533 504	24	2 283
Luxembourg	0.00	3 923	4 080	66	66
Mexico	3.65	29 983	1 071 650	34	7 264
Netherlands	1.32	3 992	184 943	20	1 041
New Zealand	0.36	4 511	48 638	263	2 411
Norway	0.53	4 064	52 816	139	1 563
Poland	2.56	4 383	534 900	75	7 517
Portugal	0.83	4 608	96 857	84	1 450
Slovak Republic	1.27	7 346	77 067	109	1 341
Spain	0.39	10 791	344 372	591	25 619
Sweden	1.44	4 624	107 104	144	3 085
Switzerland	3.41	8 420	86 491	194	893
Turkey	0.73	4 855	481 279	0	0
United Kingdom	3.36	9 535	698 579	270	15 062
United States	0.00	5 456	3 147 089	534	246 991
Partner countries					
Brazil	0.00	4 452	1 952 253	5	2 142
Hong Kong-China	0.83	4 478	72 484	8	103
Indonesia	0.31	10 761	1 971 476	0	0
Latvia	3.82	4 627	33 643	44	380
Liechtenstein	0.00	332	338	5	5
Macao-China	0.00	1 250	6 546	4	13
Russian Federation	0.99	5 974	2 153 373	35	14 716
Serbia	5.32	4 405	68 596	15	241
Thailand	0.98	5 236	637 076	5	563
Tunisia	0.34	4 721	150 875	1	31
Uruguay	0.15	5 835	33 775	18	80

Table A3.1 *(continued)*
PISA target populations and samples

		Population and sample information		Coverage indices		
		(11) Within-school exclusion rate (%)	**(12)** Overall exclusion rate (%)	**(13)** Coverage index 1: Coverage of national desired population	**(14)** Coverage index 2: Coverage of national enrolled population	**(15)** Coverage index 3: Percentage of enrolled population
OECD countries	Australia	1.51	2.15	0.98	0.97	0.93
	Austria	1.26	1.62	0.98	0.98	0.94
	Belgium	1.06	1.53	0.98	0.98	0.98
	Canada	5.26	6.83	0.93	0.93	1.00
	Czech Republic	0.18	1.20	0.99	0.99	0.97
	Denmark	4.29	5.33	0.95	0.95	0.98
	Finland	1.24	3.38	0.97	0.97	1.00
	France	1.10	3.40	0.97	0.93	1.00
	Germany	1.29	1.89	0.98	0.98	0.96
	Greece	2.46	3.19	0.97	0.97	0.97
	Hungary	0.99	3.94	0.96	0.96	0.96
	Iceland	1.97	2.59	0.97	0.97	0.99
	Ireland	2.87	4.29	0.96	0.96	0.96
	Italy	1.39	1.88	0.98	0.98	1.02
	Japan	0.00	1.02	0.99	0.99	0.97
	Korea	0.43	0.87	0.99	0.99	1.00
	Luxembourg	1.59	1.59	0.98	0.98	1.00
	Mexico	0.67	4.30	0.96	0.96	0.58
	Netherlands	0.56	1.87	0.98	0.98	1.00
	New Zealand	4.72	5.07	0.95	0.95	0.96
	Norway	2.87	3.39	0.97	0.96	0.99
	Poland	1.39	3.91	0.96	0.96	0.97
	Portugal	1.47	2.30	0.98	0.98	0.91
	Slovak Republic	1.71	2.96	0.97	0.97	0.97
	Spain	6.92	7.29	0.93	0.93	0.92
	Sweden	2.80	4.20	0.96	0.96	1.03
	Switzerland	1.02	4.39	0.96	0.96	0.97
	Turkey	0.00	0.73	0.99	0.99	0.54
	United Kingdom	2.11	5.40	0.95	0.95	0.96
	United States	7.28	7.28	0.93	0.93	1.00
Partner countries	Brazil	0.11	0.11	1.00	0.99	0.65
	Hong Kong-China	0.14	0.97	0.99	0.99	0.97
	Indonesia	0.00	0.31	1.00	0.95	0.73
	Latvia	1.12	4.89	0.95	0.95	0.99
	Liechtenstein	1.46	1.46	0.99	0.99	0.87
	Macao-China	0.20	0.20	1.00	1.00	0.83
	Russian Federation	0.68	1.66	0.98	0.98	0.95
	Serbia	0.35	5.66	0.94	0.94	0.94
	Thailand	0.09	1.06	0.99	0.99	0.84
	Tunisia	0.02	0.36	1.00	1.00	1.00
	Uruguay	0.24	0.38	1.00	1.00	0.74

Note: For a full explanation of the details in this table please refer to the *PISA 2003 Technical Report* (OECD, forthcoming).

The sampling standards used in PISA permitted countries to exclude up to a total of 5 per cent of the relevant population either by excluding schools or by excluding students within schools. All but seven countries, New Zealand (5.1 per cent), Denmark (5.3 per cent), the United Kingdom (5.4 per cent), Serbia (5.7 per cent),[1] Canada (6.8 per cent), the United States (7.3 per cent) and Spain (7.3 per cent) achieved this standard and in 20 countries the overall exclusion rate was less than 2 per cent. In some of the countries with exclusion rates exceeding 5 per cent, exclusions were inevitable. For example, in New Zealand 2.3 per cent of the students were excluded because they had less than one year of instruction in English, often because they were foreign fee-paying students and were therefore not able to follow the instructions of the assessment. When language exclusions are accounted for (*i.e.* removed from the overall exclusion rate), Denmark and New Zealand no longer had exclusion rates greater than 5 per cent. For details, see *www.pisa.oecd.org*.

Exclusions within the above limits include:

- *At the school level: i)* schools which were geographically inaccessible or where the administration of the PISA assessment was not considered feasible; and *ii)* schools that provided teaching only for students in the categories defined under "within-school exclusions", such as schools for the blind. The percentage of 15-year-olds enrolled in such schools had to be less than 2.5 per cent of the nationally desired target population (0.5% maximum for *i)* and 2% maximum for *ii)*). The magnitude, nature and justification of school-level exclusions are documented in the *PISA 2003 Technical Report* (OECD, forthcoming).
- *At the student level: i)* students with an intellectual disability; *ii)* students with a functional disability; and *iii)* students with a limited assessment language proficiency. Students could not be excluded solely because of low proficiency or normal discipline problems. The percentage of 15-year-olds excluded within schools had to be less than 2.5 per cent of the nationally desired target population.

1. For the country Serbia and Montenegro, data for Montenegro are not available. The latter accounts for 7.9 per cent of the national population. The name "Serbia" is used as a shorthand for the Serbian part of Serbia and Montenegro.

Table A3.1 describes the target population of the countries participating in PISA 2003. Further information on the target population and the implementation of PISA sampling standards can be found in the *PISA 2003 Technical Report* (OECD, forthcoming).

- **Column 1** shows the total number of 15-year-olds according to the most recent available information, which in most countries meant the year 2002 as the year before the assessment.
- **Column 2** shows the number of 15-year-olds enrolled in schools in grades 7 or above (as defined above), which is referred to as the ***eligible population.***
- **Column 3** shows the ***national desired target population***. Countries were allowed to exclude up to 0.5 per cent of students *a priori* from the eligible population, essentially for practical reasons. The following *a priori* exclusions exceed this limit but were agreed with the PISA Consortium: Australia excluded 1.04 per cent of its populations from TAFE colleges; France excluded 4.15 per cent of its students in Territoires d'Outre-Mer because they were students in outlying territories not subject to the national education system (students from outlying *departments* were included), as well as eligible students in hospitals or trade chambers; and Indonesia excluded 4.65 per cent of its students from four provinces because of security reasons.
- **Column 4** shows the number of students enrolled in schools that were excluded from the national desired target population either from the sampling frame or later in the field during data collection.
- **Column 5** shows the size of the national desired target population after subtracting the students enrolled in excluded schools. This is obtained by subtracting column 4 from column 3.
- **Column 6** shows the percentage of students enrolled in excluded schools. This is obtained by dividing column 4 by column 3 and multiplying by 100.
- **Column 7** shows the ***number of students participating in PISA 2003***. Note that this number does not account for 15-year-olds assessed as part of additional national options.
- **Column 8** shows the ***weighted number of participating students***, *i.e.* the number of students in the nationally defined target population that the PISA sample represents.
- Each country attempted to maximise the coverage of PISA's target population within the sampled schools. In the case of each sampled school, all eligible students, namely those 15 years of age, regardless of grade, were first listed. Sampled students who were to be excluded had still to be included in the sampling documentation, and a list drawn up stating the reason for their exclusion. **Column 9** indicates the total number of ***excluded students,*** which is further described and classified into specific categories in Table A3.2. **Column 10** indicates the ***weighted number of excluded students,*** *i.e.* the overall number of students in the nationally defined target population represented by the number of students excluded from the sample, which is also described and classified by exclusion categories in Table A3.2. Excluded students were excluded based on four categories: *i)* students with an intellectual disability – student has a mental or emotional disability and is cognitively delayed such that he/she cannot perform in the PISA testing situation; *ii)* students with a functional disability – student has a moderate to severe permanent physical disability such that he/she cannot perform in the PISA testing situation; and *iii)* students with a limited assessment language proficiency – student is unable to read or speak any of the languages of the assessment in the country and would be unable to overcome the language barrier in the testing situation. Typically a student who has received less than one year of instruction in the languages of the assessment may be excluded; and *iv)* other – which is a category defined by the national centres and approved by the international centre.
- **Column 11** shows the ***percentage of students excluded within schools.*** This is calculated as the weighted number of excluded students (column 10) divided by the weighted number of excluded and participating students (column 8 plus column 10).
- **Column 12** shows the ***overall exclusion rate*** which represents the weighted percentage of the national desired target population excluded from PISA either through school-level exclusions or through the exclusion of students within schools. It is calculated as the school-level exclusion rate (column 6 divided by 100) plus within-school exclusion rate (column 11 divided by 100) multiplied by 1 minus the school-level exclusion rate (column 6 divided by 100). This result is then multiplied by 100. Seven countries, namely Canada, Denmark, New Zealand, Spain, the United Kingdom, the United States and the partner country Serbia, had exclusion rates higher than 5 per cent (see also *www.oecd.org* for further information on these exclusions). When language exclusions were accounted for (*i.e.* removed from the overall exclusion rate), Denmark and New Zealand no longer had exclusion rates greater than 5 per cent.
- **Column 13** presents an index of ***the extent to which the national desired target population is covered by the PISA sample.*** Canada, Spain, the United States and the partner country Serbia were the only countries where the coverage is below 95 per cent.

Table A3.2
Exclusions

	Student exclusions (unweighted)					Student exclusions (weighted)				
	(1) Number of excluded students with disabilities (code 1)	(2) Number of excluded students with disabilities (code 2)	(3) Number of excluded students because of language (code 3)	(4) Number of excluded students for other reasons (code 4)	(5) Total number of excluded students	(6) Weighted number of excluded students with disabilities (code 1)	(7) Weighted number of excluded students with disabilities (code 2)	(8) Weighted number of excluded students because of language (code 3)	(9) Weighted number of excluded students for other reasons (code 4)	(10) Total weighted number of excluded students
OECD countries										
Australia	33	133	62	0	228	457	2 443	712	0	3 612
Austria	3	27	30	0	60	62	573	465	0	1 099
Belgium	4	49	49	0	102	64	507	622	0	1 193
Canada	100	1 590	303	0	1 993	874	13 720	3 734	0	18 328
Czech Republic	5	14	2	1	22	106	35	66	11	218
Denmark	9	70	79	56	214	101	768	861	591	2 321
Finland	15	37	20	7	79	138	334	200	53	725
France	9	31	11	0	51	1 227	5 110	1 821	0	8 158
Germany	4	21	30	6	61	768	4 526	5 347	893	11 533
Greece	14	30	31	69	144	289	555	498	1 310	2 652
Hungary	0	55	7	0	62	0	928	138	0	1 065
Iceland	12	45	22	0	79	12	45	22	0	79
Ireland	14	78	16	31	139	152	906	183	377	1 619
Italy	20	99	69	0	188	619	3 655	2 521	0	6 794
Japan	0	0	0	0	0	0	0	0	0	0
Korea	3	21	0	0	24	284	1 999	0	0	2 283
Luxembourg	2	15	45	4	66	2	15	45	4	66
Mexico	7	10	17	0	34	167	1 618	5 479	0	7 264
Netherlands	2	17	1	0	20	154	773	114	0	1 041
New Zealand	29	94	140	0	263	260	880	1 271	0	2 411
Norway	7	90	42	0	139	77	1 019	468	0	1 563
Poland	9	26	3	37	75	894	2 623	310	3 691	7 517
Portugal	14	55	15	0	84	255	929	265	0	1 450
Slovak Republic	16	74	19	0	109	108	913	320	0	1 341
Spain	34	421	136	0	591	1 594	17 246	6 779	0	25 619
Sweden	1	110	33	0	144	18	2 297	769	0	3 085
Switzerland	26	93	75	0	194	127	344	422	0	893
Turkey	0	0	0	0	0	0	0	0	0	0
United Kingdom	23	208	39	0	270	1 146	12 401	1 515	0	15 062
United States	32	431	71	0	534	14 239	201 562	31 190	0	246 991
Partner countries										
Brazil	4	1	0	0	5	1 642	500	0	0	2 142
Hong Kong-China	2	5	1	0	8	26	63	14	0	103
Indonesia	0	0	0	0	0	0	0	0	0	0
Latvia	21	23	0	0	44	148	231	0	0	380
Liechtenstein	1	0	4	0	5	1	0	4	0	5
Macao-China	4	0	0	0	4	13	0	0	0	13
Russian Federation	13	19	3	0	35	4 538	8 969	1 209	0	14 716
Serbia	5	8	2	0	15	78	129	34	0	241
Thailand	4	1	0	0	5	463	100	0	0	563
Tunisia	0	0	1	0	1	0	0	31	0	31
Uruguay	5	9	4	0	18	30	38	12	0	80

Exclusion codes:

Code 1: *Functional disability* – student has a moderate to severe permanent physical disability.

Code 2: *Intellectual disability* – student has a mental or emotional disability and has either been tested as cognitively delayed or is considered in the professional opinion of qualified staff to be cognitively delayed.

Code 3: *Limited assessment language proficiency* – student is not a native speaker of any of the languages of the assessment in the country and has limited proficiency in these languages.

Code 4: *Other* – defined by the national centres and approved by the international centre.

Note: For a full explanation of other details in this table please refer to the PISA 2003 Technical Report (OECD, forthcoming).

- **Column 14** presents an index of ***the extent to which 15-year-olds enrolled in schools are covered by the PISA sample.*** The index measures the overall proportion of the national enrolled population that is covered by the non-excluded portion of the student sample. The index takes into account both school-level and student-level exclusions. Values close to 100 indicate that the PISA sample represents the entire education system as defined for PISA 2003. The index is the weighted number of participating students (column 8) divided by the weighted number of participating and excluded students (column 8 plus column 10), times the nationally defined target population (column 5) divided by the eligible population (column 2) (times 100). The same countries with index 1 below 0.95 also had index 2 below 0.95. In addition, France also had this index below 95 per cent due to the exclusion of Territoires d'Outre Mer. This was consistent with the results from PISA 2000.
- **Column 15** presents an index of ***the percentage of enrolled population.*** This index is the total enrolled population of 15-year-olds (column 2) divided by the total population of 15-year-old students (column 1).

This high level of coverage contributes to the comparability of the assessment results. For example, even assuming that the excluded students would have systematically scored worse than those who participated, and that this relationship is moderately strong, an exclusion rate in the order of 5 per cent would likely lead to an overestimation of national mean scores of less than 5 score points (on a scale with an international mean of 500 score points and a standard deviation of 100 score points). This assessment is based on the following calculations: If the correlation between the propensity of exclusions and student performance is 0.3, resulting mean scores would likely be overestimated by 1 score point if the exclusion rate is 1 per cent, by 3 score points if the exclusion rate is 5 per cent, and by 6 score points if the exclusion rate is 10 per cent. If the correlation between the propensity of exclusions and student performance is 0.5, resulting mean scores would be overestimated by 1 score point if the exclusion rate is 1 per cent, by 5 score points if the exclusion rate is 5 per cent, and by 10 score points if the exclusion rate is 10 per cent. For this calculation, a model was employed that assumes a bivariate normal distribution for the propensity to participate and performance. For details see the *PISA 2003 Technical Report* (OECD, forthcoming).

Sampling procedures and response rates

The accuracy of any survey results depends on the quality of the information on which national samples are based as well as on the sampling procedures. Quality standards, procedures, instruments and verification mechanisms were developed for PISA that ensured that national samples yielded comparable data and that the results could be compared with confidence.

Most PISA samples were designed as two-stage stratified samples (where countries applied different sampling designs, these are documented in the *PISA 2003 Technical Report* (OECD, forthcoming)). The first stage consisted of sampling individual schools in which 15-year-old students could be enrolled. Schools were sampled systematically with probabilities proportional to size, the measure of size being a function of the estimated number of eligible (15-year-old) students enrolled. A minimum of 150 schools were selected in each country (where this number existed), although the requirements for national analyses often required a somewhat larger sample. As the schools were sampled, replacement schools were simultaneously identified, in case a sampled school chose not to participate in PISA 2003.

In the case of Iceland, Liechtenstein and Luxembourg, all schools and all eligible students within schools were included in the sample. However, since not all students in the PISA samples were assessed in all domains, these national samples represent a complete census only in respect of the assessment of mathematical literacy as the major domain, and not for the assessment of problem solving.

Experts from the PISA Consortium performed the sample selection process for each participating country and monitored it closely in those countries where they selected their own samples.

The second stage of the selection process sampled students within sampled schools. Once schools were selected, a list of each sampled school's 15-year-old students was prepared. From this list, 35 students were then selected with equal probability (all 15-year-old students were selected if fewer than 35 were enrolled).

Data quality standards in PISA required minimum participation rates for schools as well as for students. These standards were established to minimise the potential for response biases. In the case of countries meeting these standards, it was likely that any bias resulting from non-response would be negligible, *i.e.* typically smaller than the sampling error.

A minimum response rate of 85 per cent was required for the schools initially selected. Where the initial response rate of schools was between 65 and 85 per cent, however, an acceptable school response rate could still be achieved through the use of replacement schools. This procedure brought with it a risk of increased response bias. Participating countries were, therefore, encouraged to persuade as many of the schools in the original sample as possible to participate. Schools with a student participation rate between 25 and 50 per cent were not regarded as participating schools, but data from these schools were included in the database and contributed to the various estimations. Data from schools with a student participation rate of less than 25 per cent were excluded from the database.

PISA 2003 also required a minimum participation rate of 80 per cent of students within participating schools. This minimum participation rate had to be met at the national level, not necessarily by each participating school. Follow-up sessions were required in schools in which too few students had participated in the original assessment sessions. Student participation rates were calculated over all original schools, and also over all schools whether original sample or replacement schools, and from the participation of students in both the original assessment and any follow-up sessions. A student who participated in the original or follow-up cognitive sessions was regarded as a participant. Those who attended only the questionnaire session were

included in the international database and contributed to the statistics presented in this publication if he or she provided at least a description of his or her father's or mother's occupation.

Table A3.3 shows the response rates for students and schools, before and after replacement.

- **Column 1** shows the ***weighted participation rate of schools before replacement***. This is obtained by dividing column 2 by column 3.
- **Column 2** shows the ***weighted number of responding schools before school replacement*** (weighted by student enrolment).
- **Column 3** shows the ***weighted number of sampled schools before school replacement*** (including both responding and non responding schools) (weighted by student enrolment).
- **Column 4** shows the ***unweighted number of responding schools before school replacement***.
- **Column 5** shows the ***unweighted number of responding and non responding schools before school replacement***.
- **Column 6** shows the ***weighted participation rate of schools after replacement.*** This is obtained by dividing column 7 by column 8. Canada, the United Kingdom and the United States did not meet PISA's requirements for response rates before replacement, which was 85 per cent. The participation rate of Canada before replacement was 79.9 per cent (column 1) reaching 84.4 per cent after replacement, thus short by 3.1 per cent of the required 87.5 per cent. In the United Kingdom, the response rate before replacement was 64.3 (column 1) falling short of the minimum requirement by 0.7 per cent. After replacement, the participation rate increased to 77.4, still short of the final requirement. The United States achieved an initial participation rate of 64.9 before replacement reaching 68.1 after replacement
- **Column 7** shows the ***weighted number of responding schools after school replacement*** (weighted by student enrolment).
- **Column 8** shows the ***weighted number of schools sampled after school replacement*** (including both responding and nonresponding schools) (weighted by student enrolment).
- **Column 9** shows the ***unweighted number of responding schools after school replacement***.
- **Column 10** shows the ***unweighted number of responding and non responding schools after school replacement.***
- **Column 11** shows the ***weighted student participation rate after replacement.*** This is obtained by dividing column 12 by column 13. The United Kingdom was the only country where the student participation rate of 77.9 per cent was below the required 80 per cent.
- **Column 12** shows the ***weighted number of students assessed***.
- **Column 13** shows the ***weighted number of students sampled*** (including both students that were assessed and students who were absent on the day of the assessment).
- **Column 14** shows the ***unweighted number of students assessed***. Note that any students in schools with student response rates less than 50 per cent were not included in these rates (both weighted and unweighted).
- **Column 15** shows the ***unweighted number of students sampled*** (including both students that were assessed and students who were absent on the day of the assessment). Note that any students in schools with student response rates less than 50 per cent were not included in these rates (both weighted and unweighted).

Reporting of data for the United Kingdom in PISA 2003

In order to ensure that PISA yields reliable and internationally comparable data, OECD Member countries agreed on a process for the validation of all national data submissions. As the basis for this process, PISA established technical standards for the quality of datasets which countries must meet in order to be reported in OECD publications. These standards are described in detail in the *PISA 2003 Technical Report* (OECD, forthcoming). One of the requirements is that initial response rates should be 85 per cent at the school level and 80 per cent at the student level. The response rates are reported in Table A3.3.

The United Kingdom fell significantly short of these standards, with a weighted school participation rate before replacement of 64.3 per cent at the school level. As mentioned above, the Technical Standards include an approved procedure through which countries with an initial school-level response rate of at least 65 per cent could improve response rates through the use of designated replacement schools. For the United Kingdom, a school-level response rate of 96 per cent was required, but only 77.4 per cent was achieved after replacement and it was 77.9 per cent at the student level.

Table A3.3
Response rates

		Initial sample – before school replacement					Final sample – after school replacement				
		(1) Weighted school participation rate before replacement (%)	(2) Number of responding schools (weighted by enrolment)	(3) Weighted number of schools sampled (responding and non-responding) (weighted by enrolment)	(4) Number of responding schools (unweighted)	(5) Number of responding and non-responding schools (unweighted)	(6) Weighted school participation rate after replacement (%)	(7) Weighted number of responding schools (weighted by enrolment)	(8) Number of schools sampled (responing and non-responding) (weighted by enrolment)	(9) Number of responding schools (unweighted)	(10) Number of responding and non-responding schools (unweighted)
OECD countries	Australia	86.31	237 525	275 208	301	355	90.43	248 876	275 208	314	355
	Austria	99.29	87 169	87 795	192	194	99.29	87 169	87 795	192	194
	Belgium	83.40	98 423	118 010	248	296	95.63	112 775	117 924	282	296
	Canada	79.95	300 328	375 622	1 040	1 162	84.38	316 977	375 638	1 066	1 162
	Czech Republic	91.38	113 178	123 855	239	262	99.05	122 629	123 811	259	262
	Denmark	84.60	47 573	56 234	175	210	98.32	55 271	56 213	205	210
	Finland	97.39	58 209	59 766	193	197	100.00	59 766	59 766	197	197
	France	88.65	671 417	757 355	162	183	89.24	675 840	757 355	163	183
	Germany	98.06	886 841	904 387	211	216	98.82	893 879	904 559	213	216
	Greece	80.60	82 526	102 384	145	179	95.77	104 859	109 490	171	179
	Hungary	97.32	115 041	118 207	248	262	99.37	117 269	118 012	252	262
	Iceland	99.90	4 082	4 086	129	131	99.90	4 082	4 086	129	131
	Ireland	90.24	52 791	58 499	139	154	92.84	54 310	58 499	143	154
	Italy	97.54	549 168	563 039	398	406	100.00	563 039	563 039	406	406
	Japan	87.12	1 144 942	1 314 227	131	150	95.91	1 260 428	1 314 227	144	150
	Korea	95.89	589 540	614 825	143	149	100.00	614 825	614 825	149	149
	Luxembourg	99.93	4 087	4 090	29	32	99.93	4 087	4 090	29	32
	Mexico	93.98	1 132 315	1 204 851	1 090	1 154	95.45	1 150 023	1 204 851	1 102	1 154
	Netherlands	82.61	161 682	195 725	144	175	87.86	171 955	195 725	153	175
	New Zealand	91.09	48 401	53 135	158	175	97.55	51 842	53 145	171	175
	Norway	87.87	48 219	54 874	175	200	90.40	49 608	54 874	180	200
	Poland	95.12	531 479	558 752	157	166	98.09	548 168	558 853	163	166
	Portugal	99.31	106 174	106 916	152	153	99.31	106 174	106 916	152	153
	Slovak Republic	98.39	406 170	412 829	377	383	100.00	412 777	412 777	383	383
	Spain	78.92	63 629	80 626	223	284	99.08	80 394	81 141	281	284
	Sweden	99.08	112 467	113 511	185	188	99.08	112 467	113 511	185	188
	Switzerland	97.32	77 867	80 011	437	456	98.53	78 838	80 014	444	456
	Turkey	93.29	671 385	719 702	145	159	100.00	719 405	719 405	159	159
	United Kingdom	64.32	456 818	710 203	311	451	77.37	549 059	709 641	361	451
	United States	64.94	2 451 083	3 774 330	249	382	68.12	2 571 003	3 774 322	262	382
Partner countries	Brazil	93.20	2 181 287	2 340 538	213	229	99.51	2 328 972	2 340 538	228	229
	Hong Kong-China	81.89	59 216	72 312	124	151	95.90	69 345	72 312	145	151
	Indonesia	100.00	2 173 824	2 173 824	344	344	100.00	2 173 824	2 173 824	344	344
	Latvia	95.31	33 845	35 509	157	164	95.31	33 845	35 509	157	164
	Liechtenstein	100.00	348	348	12	12	100.00	348	348	12	12
	Macao-China	100.00	6 992	6 992	39	39	100.00	6 992	6 992	39	39
	Russian Federation	99.51	1 798 096	1 806 954	210	211	100.00	1 806 954	1 806 954	211	211
	Serbia	100.00	90 178	90 178	149	149	100.00	90 178	90 178	149	149
	Thailand	91.46	704 344	770 109	163	179	100.00	769 392	769 392	179	179
	Tunisia	100.00	163 555	163 555	149	149	100.00	163 555	163 555	149	149
	Uruguay	93.20	39 773	42 677	233	245	97.11	41 474	42 709	239	245

		Final sample – Students within schools after school replacement				
		(11) Weighted student participation rate after replacement (%)	(12) Number of students assessed (weighted)	(13) Number of students sampled (assessed and absent) (weighted)	(14) Number of students assessed (unweighted)	(15) Number of students sampled (assessed and absent) (unweighted)
OECD countries	Australia	83.31	176 085	211 357	12 425	15 179
	Austria	83.56	71 392	85 439	4 566	6 212
	Belgium	92.47	98 936	106 995	8 796	9 498
	Canada	83.90	233 829	278 714	27 712	31 899
	Czech Republic	89.03	106 645	119 791	6 316	7 036
	Denmark	89.88	45 356	50 464	4 216	4 687
	Finland	92.84	53 737	57 883	5 796	6 235
	France	88.11	581 957	660 491	4 214	4 774
	Germany	92.18	806 312	874 762	4 642	5 040
	Greece	95.43	96 273	100 883	4 627	4 854
	Hungary	92.87	98 996	106 594	4 764	5 132
	Iceland	85.37	3 350	3 924	3 350	3 924
	Ireland	82.58	42 009	50 873	3 852	4 670
	Italy	92.52	445 502	481 521	11 639	12 407
	Japan	95.08	1 132 200	1 190 768	4 707	4 951
	Korea	98.81	527 177	533 504	5 444	5 509
	Luxembourg	96.22	3 923	4 077	3 923	4 077
	Mexico	92.26	938 902	1 017 667	29 734	32 276
	Netherlands	88.25	144 212	163 418	3 979	4 498
	New Zealand	85.71	40 595	47 363	4 483	5 233
	Norway	87.86	41 923	47 715	4 039	4 594
	Poland	81.95	429 921	524 584	4 338	5 296
	Portugal	87.92	84 783	96 437	4 590	5 199
	Slovak Republic	90.61	312 044	344 372	10 791	11 655
	Spain	91.90	70 246	76 441	7 346	7 994
	Sweden	92.61	98 095	105 927	4 624	4 970
	Switzerland	94.70	81 026	85 556	8 415	8 880
	Turkey	96.87	466 201	481 279	4 855	5 010
	United Kingdom	77.92	419 810	538 737	9 265	11 352
	United States	82.73	1 772 279	2 142 288	5 342	6 502
Partner countries	Brazil	91.19	1 772 522	1 943 751	4 452	4 871
	Hong Kong-China	90.20	62 756	69 576	4 478	4 966
	Indonesia	98.09	1 933 839	1 971 476	10 761	10 960
	Latvia	93.88	30 043	32 001	4 627	4 940
	Liechtenstein	98.22	332	338	332	338
	Macao-China	98.02	6 642	6 775	1 250	1 274
	Russian Federation	95.71	2 061 050	2 153 373	5 974	6 253
	Serbia	91.36	62 669	68 596	4 405	4 829
	Thailand	97.81	623 093	637 076	5 236	5 339
	Tunisia	96.27	145 251	150 875	4 721	4 902
	Uruguay	90.83	29 756	32 759	5 797	6 422

The results of a subsequent bias analysis provided no evidence for any significant bias of school-level performance results but did suggest that there was potential non-response bias at student levels. The PISA Consortium concluded that it was not possible to reliably assess the magnitude, or even the direction, of this non-response bias and to correct for this. As a result, it is not possible to say with confidence that the United Kingdom's sample results reliably reflect those for the national population, with the level of accuracy required by PISA. The mean performance of the responding sample of United Kingdom pupils was 510 score points in problem solving. The uncertainties surrounding the sample and its bias are such that scores for the United Kingdom cannot reliably be compared with those of other countries.

The results are, however, accurate for many within-country comparisons between subgroups (*e.g.* males and females) and for relational analyses. The results for the United Kingdom have, therefore, been included in a separate category below the results for the other participating countries. Other data for the United Kingdom that are not reported in this volume are available at *www.pisa.oecd.org* to allow researchers to reproduce the results from the international comparisons.

All international averages and aggregate statistics include the data for the United Kingdom.

It should be noted that Scotland and Northern Ireland carried out an independent sample that met the PISA technical standards, and that these data are available at *www.pisa.oecd.org.*

Annex A4: Standard errors, significance tests and subgroup comparisons

The statistics in this report represent ***estimates*** of national performance based on samples of students rather than values that could be calculated if every student in every country had answered every question. Consequently, it is important to have measures of the degree of uncertainty of the estimates. In PISA, each estimate has an associated degree of uncertainty, which is expressed through a ***standard error***. The use of ***confidence intervals*** provides a way to make inferences about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. From an observed sample statistic it can, under the assumption of a normal distribution, be inferred that the corresponding population result would lie within the confidence interval in 95 out of 100 replications of the measurement on different samples drawn from the same population.

In many cases, readers are primarily interested in whether a given value in a particular country is different from a second value in the same or another country, *e.g.* whether females in a country perform better than males in the same country. In the tables and charts used in this report, differences are labelled as ***statistically significant*** when a difference of that size, smaller or larger, would be observed less than 5 per cent of the time, if there was actually no difference in corresponding population values. Similarly, the risk of reporting as significant if there is, in fact, no correlation between two measures is contained at 5 per cent.

Although the probability that a particular difference will falsely be declared to be statistically significant is low (5 per cent) in each single comparison, the probability of making such an error increases when several comparisons are made simultaneously.

It is possible to make an adjustment for this which reduces to 5 per cent the maximum probability that differences will be falsely declared as statistically significant at least once among all the comparisons that are made. Such an adjustment, based on the Bonferroni method, has been incorporated into the multiple comparison charts in Chapters 2. The adjusted significance test should be used when the interest of readers is to compare a country's performance with that of all other countries. For comparing a single country with another single country, no adjustment is needed.

For all other tables and charts readers should note that, if there were no real differences on a given measure, then the ***multiple comparison*** in conjunction with a 5 per cent significance level, would erroneously identify differences on 0.05 times the number of comparisons made, occasions. For example, even though the significance tests applied in PISA for identifying gender differences ensure that, for each country, the likelihood of identifying a gender difference erroneously is less than 5 per cent, a comparison showing differences for 30 countries would, on average, identify 1.35 cases (0.05 times 30) with significant gender differences, even if there were no real gender difference in any of the countries. The same applies for other statistics for which significance tests have been undertaken in this publication, such as correlations and regression coefficients.

Throughout the report, significance tests were undertaken to assess the statistical significance of the comparisons made.

Gender differences

Gender differences in student performance or other indices were tested for statistical significance. Positive differences indicate higher scores for males while negative differences indicate higher scores for females. Differences marked in bold in the tables in Annex B1 are statistically significant at the 95 per cent confidence level. For examples, see Table 5.1, Annex B1.

Performance differences between top and bottom quartiles

Differences in average performance between the top quarter and the bottom quarter on the PISA indices were tested for statistical significance. Figures marked in bold indicate that performance between the top and bottom quarter of students on the respective index is statistically significantly different at the 95 per cent confidence level.

Change in the performance per unit of the index

For many tables in Annex B1, the difference in student performance per unit of the index shown was calculated. Figures in bold indicate that the differences are statistically significantly different from zero at the 95 per cent confidence level.

Annex A5: Quality assurance

Quality assurance procedures were implemented in all parts of PISA.

The consistent quality and linguistic equivalence of the PISA assessment instruments were facilitated by providing countries with equivalent source versions of the assessment instruments in English and French and requiring countries (other than those assessing students in English and French) to prepare and consolidate two independent translations using both source versions. Precise translation and adaptation guidelines were supplied, also including instructions for the selection and training of the translators. For each country, the translation and format of the assessment instruments (including test materials, marking guides, questionnaires and manuals) were verified by expert translators appointed by the PISA Consortium (whose mother tongue was the language of instruction in the country concerned and who were knowledgeable about education systems) before they were used in the PISA Field Trial and Main Study. For further information on the PISA translation procedures see the *PISA 2003 Technical Report* (OECD, forthcoming).

The survey was implemented through standardised procedures. The PISA Consortium provided comprehensive manuals that explained the implementation of the survey, including precise instructions for the work of School Co-ordinators and scripts for Test Administrators for use during the assessment sessions. Proposed adaptations to survey procedures, or proposed modifications to the assessment session script, were submitted to the PISA Consortium for approval prior to verification. The PISA Consortium then verified the national translation and adaptation of these manuals.

To establish the credibility of PISA as valid and as unbiased and to encourage uniformity in the administration of the assessment sessions, Test Administrators in participating countries were selected using the following criteria: It was required that the Test Administrator not be the reading, mathematics or science instructor of any students in the sessions he or she would administer for PISA; it was recommended that the Test Administrator not be a member of the staff of any school where he or she would administer PISA; and it was considered preferable that the Test Administrator not be a member of the staff of any school in the PISA sample. Participating countries organised an in-person training session for Test Administrators.

Participating countries were required to ensure that Test Administrators worked with the School Co-ordinator to prepare the assessment session, including updating student tracking forms and identifying excluded students; no extra time was given for the cognitive items (while it was permissible to give extra time for the student questionnaire); no instrument was administered before the two 1-hour parts of the cognitive session; Test Administrators recorded the student participation status on the student tracking forms and filled in a Session Report Form; no cognitive instrument was permitted to be photocopied; no cognitive instrument could be viewed by school staff before the assessment session; and that Test Administrators returned the material to the national centre immediately after the assessment sessions.

National Project Managers were encouraged to organise a follow-up session when more than 15 per cent of the PISA sample was not able to attend the original assessment session.

National Quality Monitors from the PISA Consortium visited all national centres to review data-collection procedures. Finally, School Quality Monitors from the PISA Consortium visited a sample of 15 schools during the assessment. For further information on the field operations see the *PISA 2003 Technical Report* (OECD, forthcoming).

Marking procedures were designed to ensure consistent and accurate application of the marking guides outlined in the PISA Operations manuals. National Project Managers were required to submit proposed modifications to these procedures to the Consortium for approval. Reliability studies to analyse the consistency of marking were implemented, these are discussed in more detail below.

Software specially designed for PISA 2003 facilitated data entry, detected common errors during data entry, and facilitated the process of data cleaning. Training sessions familiarised National Project Managers with these procedures.

For a description of the quality assurance procedures applied in PISA and the results see the *PISA 2003 Technical Report* (OECD, forthcoming).

Annex A6: Development of the PISA assessment instruments

The development of the PISA 2003 assessment instruments was an interactive process between the PISA Consortium, the various expert committees, the PISA Governing Board and national experts. A panel of international experts led, in close consultation with participating countries, the identification of the range of skills and competencies that were, in the respective assessment domains, considered to be crucial for an individual's capacity to fully participate in and contribute to a successful modern society. A description of the assessment domains – the assessment framework – was then used by participating countries, and other test development professionals, as they contributed assessment materials. The development of this assessment framework involved the following steps:

- development of a working definition for the domain and description of the assumptions that underlay that definition;
- evaluation of how to organise the set of tasks constructed in order to report to policy-makers and researchers on performance in each assessment domain among 15-year-old students in participating countries;
- identification of a set of key characteristics to be taken into account when assessment tasks were constructed for international use;
- operationalisation of the set of key characteristics to be used in test construction, with definitions based on existing literature and the experience of other large-scale assessments;
- validation of the variables, and assessment of the contribution which each made to the understanding of task difficulty in participating countries; and
- preparation of an interpretative scheme for the results.

The frameworks were agreed at both scientific and policy levels and subsequently provided the basis for the development of the assessment instruments. The frameworks are described in *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003b). They provided a common language and a vehicle for participating countries to develop a consensus as to the measurement goals of PISA.

Assessment items were then developed to reflect the intentions of the frameworks and were piloted in a Field Trial in all participating countries before a final set of items was selected for the PISA 2003 Main Study.

Due attention was paid to reflecting the national, cultural and linguistic variety among OECD countries. As part of this effort the PISA Consortium used professional test item development teams in several different countries, including Australia, the United Kingdom, the Netherlands and Japan. In addition to the items that were developed by the PISA Consortium teams, assessment material was contributed by participating countries. The Consortium's multi-national team of test developers deemed a substantial amount of this submitted material as appropriate given the requirements laid out by the PISA assessment frameworks. As a result, the item pool included assessment items from Argentina, Australia, Austria, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Korea, the Netherlands, New Zealand, Norway, Portugal, Sweden, Switzerland, and the United States. About one-third of items selected for inclusion in the Field Trial were submitted by participating countries, and about 37 per cent of items selected for the Main Study were from participating countries.

Approximately 232 units comprising about 530 items were included in item bundles for national review, in the mathematics, problem solving and science areas. After the first consultation process, the Field Trial included 115 mathematics units with 217 mathematics items. Of these mathematics units, the stimulus material for 53 came from national contributions, 80 originated with the PISA Consortium, and one unit came from the Third International Mathematics and Science Study (TIMSS).

Each item included in the assessment pool was then rated by each country: for potential cultural, gender or other bias; for relevance to 15-year-olds in school and non-school contexts; and for familiarity and level of interest. A first consultation of countries on the item pool was undertaken as part of the process of developing the Field Trial assessment instruments. A second consultation was undertaken after the Field Trial to assist in the final selection of items for the Main Study.

Following the Field Trial, in which all items were tested in all participating countries, test developers and expert groups considered a variety of aspects in selecting the items for the Main Study: *i)* the results from the Field Trial, *ii)* the outcome of the item review from countries, and *iii)* queries received during the Field Trial marking process. The test developers and expert groups selected a final set of items in October 2002 which, following a period of negotiation, was adopted by participating countries at both scientific and policy levels.

The Main Study included 54 mathematics units with 85 items. Twenty-four of these units originated from material submitted by participating countries. Twenty-eight of the units came from one or other of the Consortium teams, and two originated as TIMSS material. The Main Study instruments also included eight reading units (28 items), 13 science units (35 items) and ten problem-solving units (19 items).

Five item types were used in the PISA assessment instruments:

- ***Open-constructed response items:*** in these items, students constructed a longer response, allowing for the possibility of a broad range of divergent, individual responses and differing viewpoints. These items usually asked students to relate information or ideas in the stimulus text to their own experience or opinions, with the acceptability depending less on the position taken by the student than on the ability to use what they had read when justifying or explaining that position. Partial credit was often permitted for partially correct or less sophisticated answers, and all of these items were marked by hand.
- ***Closed-constructed response items:*** these items required students to construct their own responses, there being a limited range of acceptable answers. Most of these items were scored dichotomously with a few items included in the marking process.
- ***Short response items:*** as in the closed constructed-response items, students were to provide a brief answer, but there was a wide range of possible answers. These items were hand-marked, thus allowing for dichotomous as well as partial credit.
- ***Complex multiple-choice items:*** in these items, the student made a series of choices, usually binary. Students indicated their answer by circling a word or short phrase (for example *yes* or *no*) for each point. These items were scored dichotomously for each choice, yielding the possibility of full or partial credit for the whole item.
- ***Multiple-choice items:*** these items required students to circle a letter to indicate one choice among four or five alternatives, each of which might be a number, a word, a phrase or a sentence. They were scored dichotomously.

PISA 2003 was designed to yield group-level information in a broad range of content. The PISA assessment of mathematics included material allowing for a total of 210 minutes of assessment time. The reading, science and problem-solving assessments each included 60 minutes of assessment time. Each student, however, sat assessments lasting a total of 120 minutes.

In order to cover the intended broad range of content while meeting the limit of 120 minutes of individual assessment time, the assessment in each domain was divided into clusters, organised into thirteen booklets. There were seven 30-minute mathematics clusters, two 30-minute clusters for each of reading, science and problem solving. In PISA 2003, every student answered mathematics items, and over half the students answered items on reading, science and problem solving.

This assessment design was balanced so that each item cluster appeared four times, once in each of four possible locations in a booklet. Further, each cluster appeared once with each other cluster. The final design, therefore, ensured that a representative sample responded to each cluster of items.

For further information on the development of the PISA assessment instruments and the PISA assessment design, see the *PISA 2003 Technical Report* (OECD, forthcoming).

Annex A7: Reliability of the marking of open-ended items

The process of marking open-ended items was an important step in ensuring the quality and comparability of results from PISA.

Detailed guidelines contributed to a marking process that was accurate and consistent across countries. The marking guidelines consisted of marking manuals, training materials for recruiting markers, and workshop materials used for the training of national markers. Before national training, the PISA Consortium organised training sessions to present the material and train the marking co-ordinators from the participating countries. The latter were then responsible for training their national markers.

For each assessment item, the relevant marking manual described the aim of the question and how to code students' responses to each item. This description included the credit labels – full credit, partial credit or no credit – attached to the possible categories of responses. PISA 2003 also included a system of double-digit coding for the mathematics and science items in which the first digit represented the score and the second digit represented different strategies or approaches that students used to solve the problem. The second digit generated national profiles of student strategies and misconceptions. By way of illustration, the marking manuals also included real examples of students' responses (drawn from the Field Trial) accompanied by a rationale for their classification.

In each country, a sub-sample of assessment booklets was marked independently by four markers and examined by the PISA Consortium. In order to examine the consistency of this marking process in more detail within each country and to estimate the magnitude of the variance components associated with the use of markers, the PISA Consortium conducted an inter-marker reliability study on the sub-sample of assessment booklets. Homogeneity analysis was applied to the national sets of multiple marking and compared with the results of the Field Trial. For details see the *PISA 2003 Technical Report* (OECD, forthcoming).

At the between-country level, an inter-country reliability study was carried out on a sub-set of items. The aim was to check whether the marking given by national markers was of equal severity in each country, both overall and for particular items. In this process, independent marking of the original booklets was undertaken by trained multilingual staff and compared to the ratings by the national markers in the various countries. The results showed that very consistent marks were achieved across countries. The average index of agreement in the inter-country reliability study was 92 per cent (out of 71 941 student responses that were independently scored by the international verifiers). Agreement meant both cases where the international verifier agreed with at least three of the national markers and cases where the verifier disagreed with the national markers, but the adjudication undertaken by the PISA Consortium's test developers concluded, after reviewing the translated student's answer, that the national markers had given the correct mark. Only 6 countries had rates of agreement lower than 90 per cent (with a minimum of 86 per cent in Spain [Catalonian region]). On average, marking was too harsh in 1.8 per cent of cases and too lenient in 3.1 per cent of cases. The highest per cent of too harsh codes (7 per cent) was observed for the science items in Portugal, and the highest per cent of too lenient marks (10 per cent) was observed for the science items in Indonesia. A full description of this process and the results can be found in the *PISA 2003 Technical Report* (OECD, forthcoming).

Annex B

DATA TABLES

Table 2.1

Percentage of students at each level of proficiency on the problem-solving scale

		Proficiency levels							
		Below Level 1 (below 405 score points)		Level 1 (from 405 to 499 score points)		Level 2 (from 499 to 592 score points)		Level 3 (above 592 score points)	
		%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD countries	Australia	9	(0.6)	26	(0.7)	39	(0.8)	26	(0.8)
	Austria	14	(1.0)	32	(1.1)	37	(1.1)	17	(1.2)
	Belgium	14	(0.7)	24	(0.7)	34	(0.8)	28	(0.9)
	Canada	8	(0.5)	27	(0.7)	40	(0.7)	25	(0.7)
	Czech Republic	12	(1.1)	29	(1.2)	37	(1.1)	22	(1.2)
	Denmark	10	(0.8)	30	(0.9)	39	(0.9)	20	(0.9)
	Finland	5	(0.5)	22	(0.8)	43	(0.8)	30	(0.9)
	France	12	(1.0)	28	(1.0)	37	(1.1)	23	(1.0)
	Germany	14	(1.0)	28	(1.1)	36	(1.5)	22	(1.4)
	Greece	33	(1.5)	36	(1.0)	24	(1.2)	7	(0.8)
	Hungary	16	(1.0)	32	(1.4)	35	(1.2)	17	(1.2)
	Iceland	12	(0.7)	32	(1.0)	40	(1.0)	15	(0.6)
	Ireland	13	(0.9)	37	(1.2)	38	(1.0)	12	(0.8)
	Italy	25	(1.3)	35	(1.2)	30	(1.0)	11	(0.7)
	Japan	10	(1.0)	20	(1.0)	34	(1.2)	36	(1.6)
	Korea	5	(0.5)	22	(1.0)	41	(1.1)	32	(1.3)
	Luxembourg	17	(0.7)	34	(1.0)	35	(1.0)	14	(0.6)
	Mexico	58	(1.9)	30	(1.1)	11	(1.0)	1	(0.2)
	Netherlands	11	(1.1)	30	(1.3)	36	(1.4)	23	(1.1)
	New Zealand	10	(0.8)	25	(0.8)	36	(1.0)	28	(0.9)
	Norway	19	(0.9)	33	(1.2)	33	(1.0)	15	(0.8)
	Poland	18	(1.0)	37	(1.0)	34	(1.1)	12	(0.7)
	Portugal	24	(1.7)	36	(1.1)	31	(1.4)	9	(0.6)
	Slovak Republic	17	(1.4)	34	(1.2)	34	(1.3)	14	(1.0)
	Spain	20	(0.9)	35	(1.1)	33	(1.2)	12	(0.8)
	Sweden	12	(0.9)	32	(1.1)	38	(1.0)	17	(1.0)
	Switzerland	11	(0.7)	27	(1.0)	39	(1.1)	23	(1.4)
	Turkey	51	(2.5)	33	(1.6)	12	(1.6)	4	(1.2)
	United States	24	(1.1)	34	(0.8)	30	(1.0)	12	(0.8)
	OECD total	***22***	***(0.4)***	***30***	***(0.3)***	***31***	***(0.4)***	***17***	***(0.3)***
	OECD average	***17***	***(0.2)***	***30***	***(0.2)***	***34***	***(0.2)***	***18***	***(0.2)***
Partner countries	Brazil	64	(1.9)	26	(1.5)	9	(1.1)	2	(0.5)
	Hong Kong-China	8	(1.1)	21	(1.0)	36	(1.2)	35	(1.4)
	Indonesia	73	(1.7)	23	(1.4)	4	(0.6)	0	(0.1)
	Latvia	20	(1.5)	36	(1.3)	32	(1.4)	12	(1.0)
	Liechtenstein	10	(1.5)	26	(2.4)	37	(3.6)	27	(2.6)
	Macao-China	6	(0.8)	27	(1.4)	42	(2.0)	24	(1.6)
	Russian Federation	23	(1.7)	34	(1.0)	31	(1.3)	12	(1.0)
	Serbia	43	(1.7)	39	(1.2)	16	(1.2)	2	(0.3)
	Thailand	41	(1.6)	40	(1.1)	16	(1.1)	3	(0.5)
	Tunisia	77	(1.1)	20	(0.8)	2	(0.5)	0	(0.1)
	Uruguay	47	(1.6)	31	(1.3)	18	(1.2)	5	(0.5)
	United Kingdom[1]	m	m	m	m	m	m	m	m

1. Response rate too low to ensure comparability (see Annex A3).

Table 2.2
Mean score and variation in student performance on the problem-solving scale

		Mean score		Standard deviation		5th percentile		10th percentile		25th percentile		75th percentile		90th percentile		95th percentile	
		Mean	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD countries	Australia	530	(2.0)	91	(1.4)	371	(4.1)	409	(3.5)	469	(2.8)	594	(2.1)	644	(2.7)	672	(3.4)
	Austria	506	(3.2)	90	(1.7)	357	(5.1)	388	(4.5)	443	(4.1)	569	(4.0)	621	(4.2)	651	(4.6)
	Belgium	525	(2.2)	104	(1.5)	340	(5.0)	383	(4.5)	456	(3.3)	602	(2.6)	653	(2.0)	681	(2.0)
	Canada	529	(1.7)	88	(0.9)	379	(2.4)	414	(2.8)	471	(2.5)	591	(1.9)	640	(2.1)	669	(2.4)
	Czech Republic	516	(3.4)	93	(1.9)	356	(8.6)	394	(6.2)	454	(4.4)	582	(3.6)	634	(3.9)	663	(4.0)
	Denmark	517	(2.5)	87	(1.5)	369	(5.0)	402	(4.3)	459	(3.1)	578	(2.8)	627	(3.4)	655	(3.7)
	Finland	548	(1.9)	82	(1.2)	409	(4.7)	442	(2.8)	495	(2.5)	604	(2.3)	650	(2.3)	677	(3.6)
	France	519	(2.7)	93	(2.1)	358	(6.1)	396	(4.8)	459	(3.9)	586	(3.0)	635	(3.7)	662	(4.5)
	Germany	513	(3.2)	95	(1.8)	351	(5.9)	383	(5.3)	447	(4.8)	583	(4.3)	632	(2.7)	658	(3.2)
	Greece	448	(4.0)	99	(1.7)	283	(5.6)	319	(5.3)	383	(4.5)	517	(4.6)	574	(5.7)	607	(5.6)
	Hungary	501	(2.9)	94	(2.0)	343	(5.8)	378	(4.1)	436	(3.8)	567	(3.9)	622	(4.3)	653	(5.4)
	Iceland	505	(1.4)	85	(1.1)	358	(5.5)	393	(3.3)	450	(2.2)	564	(2.0)	609	(2.3)	634	(3.6)
	Ireland	498	(2.3)	80	(1.4)	364	(4.5)	395	(3.8)	445	(3.1)	555	(2.7)	601	(2.8)	625	(3.2)
	Italy	469	(3.1)	102	(2.1)	289	(8.7)	334	(6.5)	406	(4.7)	540	(3.0)	595	(3.4)	627	(3.6)
	Japan	547	(4.1)	105	(2.7)	362	(8.3)	406	(6.8)	481	(5.7)	621	(4.2)	675	(4.6)	705	(6.0)
	Korea	550	(3.1)	86	(2.0)	404	(4.6)	438	(5.2)	494	(3.9)	610	(3.5)	658	(4.2)	686	(5.5)
	Luxembourg	494	(1.4)	92	(1.0)	339	(3.7)	373	(2.3)	432	(2.4)	558	(2.2)	610	(2.6)	640	(3.4)
	Mexico	384	(4.3)	96	(2.0)	227	(5.4)	262	(5.2)	317	(5.2)	451	(5.1)	509	(5.7)	542	(6.5)
	Netherlands	520	(3.0)	89	(2.0)	372	(5.9)	401	(5.1)	456	(4.9)	587	(3.6)	636	(3.3)	662	(3.7)
	New Zealand	533	(2.2)	96	(1.2)	370	(3.8)	406	(4.2)	468	(3.7)	601	(2.4)	653	(2.5)	682	(2.8)
	Norway	490	(2.6)	99	(1.7)	322	(5.5)	361	(4.6)	424	(3.7)	559	(3.3)	615	(4.2)	645	(4.4)
	Poland	487	(2.8)	90	(1.7)	338	(5.6)	372	(4.1)	428	(3.1)	548	(3.0)	600	(3.5)	632	(4.5)
	Portugal	470	(3.9)	92	(2.1)	311	(7.9)	345	(6.8)	409	(5.7)	534	(3.6)	586	(3.5)	614	(3.5)
	Slovak Republic	492	(3.4)	93	(2.4)	337	(7.1)	370	(5.9)	430	(4.7)	558	(3.6)	609	(3.8)	638	(4.2)
	Spain	482	(2.7)	94	(1.3)	322	(4.8)	361	(4.1)	421	(3.5)	547	(3.2)	599	(3.9)	629	(3.3)
	Sweden	509	(2.4)	88	(1.6)	360	(6.4)	395	(4.4)	451	(3.0)	571	(3.1)	619	(3.8)	647	(3.6)
	Switzerland	521	(3.0)	94	(1.9)	358	(5.7)	397	(4.0)	461	(3.3)	587	(3.9)	637	(4.6)	666	(5.2)
	Turkey	408	(6.0)	97	(4.4)	257	(7.8)	291	(6.6)	343	(5.2)	466	(7.7)	531	(11.9)	577	(18.6)
	United States	477	(3.1)	98	(1.3)	312	(5.6)	347	(4.6)	410	(4.1)	548	(3.3)	604	(4.0)	635	(4.2)
	OECD total	***490***	***(1.2)***	***106***	***(0.8)***	***308***	***(2.7)***	***348***	***(2.2)***	***418***	***(1.7)***	***566***	***(1.3)***	***624***	***(1.3)***	***656***	***(1.4)***
	OECD average	***500***	***(0.6)***	***100***	***(0.4)***	***328***	***(1.7)***	***368***	***(1.3)***	***434***	***(1.1)***	***571***	***(0.8)***	***625***	***(0.8)***	***656***	***(0.8)***
Partner countries	Brazil	371	(4.8)	100	(2.6)	211	(7.5)	244	(6.1)	302	(4.7)	438	(5.7)	501	(7.3)	538	(8.3)
	Hong Kong-China	548	(4.2)	97	(2.9)	376	(10.5)	420	(7.9)	487	(6.1)	617	(3.2)	664	(2.9)	690	(3.7)
	Indonesia	361	(3.3)	73	(1.7)	245	(4.2)	270	(3.8)	312	(3.6)	409	(4.1)	457	(5.5)	487	(5.9)
	Latvia	483	(3.9)	92	(1.7)	326	(7.0)	362	(6.0)	420	(5.4)	547	(4.6)	599	(4.1)	628	(4.9)
	Liechtenstein	529	(3.9)	93	(4.2)	369	(14.9)	404	(11.1)	468	(6.0)	599	(9.3)	644	(10.5)	672	(12.0)
	Macao-China	532	(2.5)	81	(2.6)	395	(6.4)	425	(5.6)	478	(3.7)	590	(4.3)	633	(5.4)	659	(6.5)
	Russian Federation	479	(4.6)	99	(2.1)	314	(7.7)	351	(7.0)	413	(5.7)	546	(5.1)	604	(5.0)	637	(5.6)
	Serbia	420	(3.3)	86	(1.6)	279	(4.2)	311	(4.4)	363	(3.9)	478	(4.2)	530	(4.9)	560	(5.1)
	Thailand	425	(2.7)	82	(1.6)	293	(3.9)	322	(3.4)	369	(2.6)	478	(4.0)	532	(4.0)	565	(6.0)
	Tunisia	345	(2.1)	80	(1.4)	213	(4.3)	243	(3.1)	291	(2.5)	400	(2.8)	446	(4.1)	474	(5.0)
	Uruguay	411	(3.7)	112	(1.9)	224	(5.7)	265	(5.1)	334	(4.7)	488	(5.5)	552	(5.0)	589	(5.7)
	United Kingdom[1]	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m

1. Response rate too low to ensure comparability (see Annex A3).

Table 3.1
Factor loadings of mathematics, reading and problem-solving items
The strength of association of each PISA item with two different presumed factors, calculated from students' responses

Problem-solving item (shown in bold) | Mathematics item | Reading item

Item loads higher on factor 1 than on factor 2						Item loads higher on factor 2 than on factor 1		
Items	Factor 1 Loading	Factor 2 Loading	Items	Factor 1 Loading	Factor 2 Loading	Items	Factor 1 Loading	Factor 2 Loading
BRAILLE – Question 2	0.400		THE BEST CAR – Question 2	0.279	0.106	DRUGGED SPIDERS – Question 5	–0.136	0.642
HOLIDAY – Question 2	**0.393**		CASH WITHDRAWAL – Question 2	0.278		OPTICIAN – Question 6	–0.100	0.609
SKATEBOARD – Question 13	0.391		CARBON DIOXIDE – Question 2	0.276		OPTICIAN – Question 3		0.576
POPULATION PYRAMIDS – Question 3	0.380		CARPENTER – Question 1	0.275		EXCHANGE – Question 6		0.576
BICYCLES – Question 3	0.372		CARBON DIOXIDE – Question 3	0.271		EXCHANGE – Question 1	–0.110	0.557
BOOKSHELVES – Question 1	0.368		**CHILDREN'S CAMP – Question 1**	**0.271**		TELEPHONE – Question 1	–0.117	0.556
NUMBER CUBES – Question 3	0.368		SEEING THE TOWER – Question 1	0.265		DRUGGED SPIDERS – Question 3		0.541
POPULATION PYRAMIDS – Question 2	0.359		HEIGHT – Question 1	0.261	0.188	DRUGGED SPIDERS – Question 2		0.521
CONTAINERS – Question 1	0.356		CHAIR LIFT – Question 2	0.258		DRUGGED SPIDERS – Question 1	–0.119	0.521
CUBES – Question 1	0.355		CHAIR LIFT – Question 1	0.256		OPTICIAN – Question 2		0.494
SUPPORT FOR PRESIDENT – Question 1	0.353		NUMBER CHECK – Question 1	0.255		TELEPHONE – Question 5		0.487
CARBON DIOXIDE – Question 1	0.350		PIPELINES – Question 1	0.254		OPTICIAN – Question 1		0.371
LABELS – Question 1	0.348		RUNNING TRACKS – Question 1	0.252	0.171	TELEPHONE – Question 2		0.324
THERMOMETER CRICKET – Question 1	0.347		A VIEW ROOM – Question 1	0.250		SOUTH POLE – Question 5	0.102	0.291
CASH WITHDRAWAL – Question 1	0.346		RUNNING TRACKS – Question 2	0.246	0.160	SOUTH POLE – Question 2	0.125	0.279
IRRIGATION – Question 3	**0.346**		THERMOMETER CRICKET – Question 2	0.236		SOUTH POLE – Question 1	0.169	0.261
THE FENCE – Question 1	0.346		**CINEMA OUTING – Question 1**	**0.234**		SOUTH POLE – Question 4	0.106	0.257
ROBBERIES – Question 15	0.345		**DESIGN BY NUMBERS© – Question 1**	**0.234**		SOUTH POLE – Question 6		0.254
TILE ARRANGEMENT – Question 1	0.342		STEP PATTERN – Question 1	0.233		EMPLOYMENT – Question 1	0.150	0.248
BRICKS – Question 1	0.338		THIRD SIDE – Question 1	0.232		EMPLOYMENT – Question 1 (E)	0.174	0.235
SCIENCE TESTS – Question 1	0.337		**LIBRARY SYSTEM – Question 1**	**0.232**		SHIRTS – Question 7	0.106	0.234
COLOURED CANDIES – Question 1	0.336		CAR DRIVE – Question 2	0.223		EMPLOYMENT – Question 2	0.121	0.233
SKATEBOARD – Question 12	0.335		LITTER – Question 1	0.221		SHIRTS – Question 5	0.153	0.230
WALKING – Question 5	0.334		STOP THE CAR – Question 1	0.221	0.142	EXCHANGE RATE – Question 10	0.201	0.227
CAR DRIVE – Question 3	0.333		DIVING – Question 1	0.216	0.193	GROWING UP – Question 7	0.181	0.223
COURSE DESIGN – Question 1	**0.328**		CUBE PAINTING – Question 4	0.215	0.122	EXCHANGE RATE – Question 9	0.165	0.217
IRRIGATION – Question 2	**0.321**		STAIRCASE – Question 2	0.213		SHIRTS – Question 4	0.145	0.200
LIBRARY SYSTEM – Question 2	**0.318**		RUNNING TRACKS – Question 3	0.213	0.145	GROWING UP – Question 6	0.182	0.196
POPULATION PYRAMIDS – Question 1	0.315		**DESIGN BY NUMBERS© – Question 2**	**0.207**		EXCHANGE RATE – Question 11	0.193	0.193
IRRIGATION – Question 1	**0.313**		HEIGHT – Question 3	0.207		THE BEST CAR – Question 1	0.173	0.187
HOLIDAY – Question 1	**0.310**		TOSSING COINS – Question 1	0.205		AESOP – Question 1		0.179
BICYCLES – Question 2	0.308		CHOICES – Question 1	0.204		DIVING – Question 2	0.153	0.172
TRANSPORT – Question 1	0.308		CUBE PAINTING – Question 1	0.202	0.126	MAKING A BOOKLET – Question 1	0.141	0.147
EARTHQUAKE – Question 1	0.305		GROWING UP – Question 8	0.201	0.198	COMPUTER GAME – Question 1		0.128
LOTTERIES – Question 1	0.304		CUBE PAINTING – Question 3	0.197	0.169			
INTERNET RELAY CHAT – Question 2	0.303		MAP – Question 1	0.196				
ENERGY NEEDS – Question 2	**0.303**		TELEPHONE RATES – Question 1	0.194	0.133			
WALKING – Question 4	0.301		**CINEMA OUTING – Question 2**	**0.189**				
TEST SCORES – Question 16	0.298		FORECAST OF RAIN – Question 1	0.189	0.130			
SKATEBOARD – Question 14	0.298		**FREEZER – Question 1**	**0.188**				
NUMBER CHECK – Question 2	0.297		CUBE PAINTING – Question 2	0.184				
POPULATION PYRAMIDS – Question 4	0.295		EXPORTS – Question 17	0.176				
TRANSIT SYSTEM – Question 1	**0.292**		**FREEZER – Question 2**	**0.170**				
INTERNET RELAY CHAT – Question 1	0.287		HEIGHT – Question 2	0.160				
BICYCLES – Question 1	0.286		RUNNING TIME – Question 1	0.160				
DESIGN BY NUMBERS© – Question 3	**0.285**		**ENERGY NEEDS – Question 1**	**0.157**				
EXPORTS – Question 18	0.281		CAR DRIVE – Question 1	0.150				

Note: The items are arranged in order of the magnitudes of the factor loadings. Columns 1 and 2 show items where the factor 1 loading is higher than the factor 2 loading for an item, and the items are arranged in descending order of the factor 1 loadings. Column 3 shows items where the factor 2 loading is higher than the factor 1 loading for an item, and the items are arranged in descending order of factor 2 loadings.
Source: OECD, PISA 2003 database.

Table 3.2
Difference between mean scores in mathematics and problem solving

		(Math. – P.S.)	S.E.
OECD countries	Australia	**–5.6**	(0.70)
	Austria	–0.5	(0.78)
	Belgium	**4.0**	(0.78)
	Canada	**3.2**	(0.53)
	Czech Republic	0.0	(0.85)
	Denmark	–2.5	(1.18)
	Finland	**–3.3**	(0.68)
	France	**–8.4**	(1.01)
	Germany	**–10.5**	(0.91)
	Greece	–3.6	(1.33)
	Hungary	**–11.1**	(0.82)
	Iceland	**10.5**	(0.98)
	Ireland	**4.4**	(0.79)
	Italy	–3.8	(1.22)
	Japan	**–13.1**	(1.06)
	Korea	**–8.2**	(0.79)
	Luxembourg	–0.4	(1.13)
	Mexico	0.8	(1.30)
	Netherlands	**17.6**	(1.33)
	New Zealand	**–9.3**	(1.01)
	Norway	**5.4**	(1.24)
	Poland	3.7	(1.20)
	Portugal	**–3.8**	(0.88)
	Slovak Republic	**6.4**	(0.99)
	Spain	2.8	(0.97)
	Sweden	0.5	(1.27)
	Switzerland	**5.3**	(0.83)
	Turkey	**15.9**	(1.27)
	United States	**5.5**	(0.62)
Partner countries	Brazil	**–15.0**	(0.99)
	Hong Kong-China	2.5	(1.20)
	Indonesia	–1.3	(1.22)
	Latvia	0.9	(1.33)
	Liechtenstein	6.4	(2.80)
	Macao-China	–5.2	(1.77)
	Russian Federation	**–10.2**	(1.35)
	Serbia	**16.7**	(1.30)
	Thailand	**–7.9**	(1.23)
	Tunisia	**14.0**	(1.23)
	Uruguay	**11.6**	(1.41)
	United Kingdom[1]	m	m

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.1
Gender differences in mean score in student performance on the problem-solving, mathematics and reading scales and percentage of males and females below Level 1 and at Level 3 of the problem-solving scale

	Problem solving											
	Males				Females						Explained variance in student performance (r-squared × 100)	
	Mean score		Standard deviation		Mean score		Standard deviation		Gender difference			
	Mean	S.E.	S.D.	S.E.	Mean	S.E.	S.D.	S.E.	Score dif.	S.E.	%	S.E.
OECD countries												
Australia	527	(2.7)	95	(1.5)	533	(2.5)	87	(1.8)	–6.40	(3.3)	0.1	(0.1)
Austria	505	(3.9)	94	(2.1)	508	(3.8)	86	(2.3)	–2.88	(4.3)	0.0	(0.1)
Belgium	522	(3.1)	108	(1.9)	527	(3.2)	101	(2.2)	–3.49	(4.5)	0.0	(0.1)
Canada	533	(2.0)	93	(1.1)	532	(1.8)	84	(1.2)	0.49	(2.1)	0.0	(0.0)
Czech Republic	520	(4.1)	95	(2.1)	513	(4.3)	91	(2.6)	6.53	(5.0)	0.1	(0.2)
Denmark	519	(3.1)	88	(1.8)	514	(2.9)	86	(2.2)	4.90	(3.2)	0.1	(0.1)
Finland	543	(2.5)	87	(1.5)	553	(2.2)	77	(1.4)	–9.99	(3.0)	0.4	(0.2)
France	519	(3.8)	97	(2.9)	520	(2.9)	89	(2.1)	–0.78	(4.1)	0.0	(0.0)
Germany	511	(3.9)	96	(2.3)	517	(3.7)	92	(2.2)	–5.71	(3.9)	0.1	(0.1)
Greece	450	(4.9)	104	(2.3)	448	(4.1)	94	(1.9)	1.94	(4.4)	0.0	(0.1)
Hungary	499	(3.4)	96	(2.3)	503	(3.4)	92	(2.6)	–3.71	(3.7)	0.0	(0.1)
Iceland	490	(2.2)	89	(1.5)	520	(2.5)	78	(1.6)	**–30.46**	(3.9)	3.2	(0.8)
Ireland	499	(2.8)	81	(1.9)	498	(3.5)	78	(1.7)	0.52	(4.2)	0.0	(0.1)
Italy	467	(5.0)	110	(3.1)	471	(3.5)	94	(2.0)	–4.06	(6.0)	0.0	(0.1)
Japan	546	(5.7)	111	(3.6)	548	(4.1)	99	(3.0)	–2.41	(5.7)	0.0	(0.1)
Korea	554	(4.0)	88	(2.3)	546	(4.8)	84	(2.6)	8.15	(6.1)	0.2	(0.3)
Luxembourg	495	(2.4)	95	(1.6)	493	(1.9)	88	(1.6)	2.37	(3.3)	0.0	(0.1)
Mexico	387	(5.0)	97	(2.6)	382	(4.7)	95	(2.6)	5.08	(4.5)	0.1	(0.1)
Netherlands	522	(3.6)	89	(2.2)	518	(3.6)	89	(2.6)	4.45	(4.1)	0.1	(0.1)
New Zealand	531	(2.6)	99	(1.6)	534	(3.1)	92	(1.7)	–3.27	(3.8)	0.0	(0.1)
Norway	486	(3.1)	103	(2.0)	494	(3.2)	94	(2.7)	**–8.46**	(3.6)	0.2	(0.2)
Poland	486	(3.4)	96	(2.4)	487	(3.0)	84	(1.8)	–1.07	(3.1)	0.0	(0.0)
Portugal	470	(4.6)	99	(2.5)	470	(3.9)	86	(2.1)	0.01	(3.5)	0.0	(0.0)
Slovak Republic	495	(4.1)	95	(2.8)	488	(3.6)	90	(2.6)	6.93	(3.7)	0.1	(0.2)
Spain	479	(3.6)	98	(1.8)	485	(2.6)	89	(1.5)	–6.04	(3.1)	0.1	(0.1)
Sweden	504	(3.0)	90	(2.0)	514	(2.8)	86	(2.1)	**–9.90**	(3.1)	0.3	(0.2)
Switzerland	520	(4.0)	96	(2.6)	523	(3.3)	92	(1.8)	–2.46	(4.1)	0.0	(0.1)
Turkey	408	(7.3)	102	(5.2)	406	(5.8)	89	(4.1)	2.01	(5.8)	0.0	(0.1)
United States	477	(3.4)	101	(1.6)	478	(3.5)	95	(1.9)	–0.95	(3.0)	0.0	(0.0)
OECD total	*489*	*(1.4)*	*109*	*(1.1)*	*490*	*(1.3)*	*103*	*(0.8)*	*–0.67*	*(1.5)*	*0.0*	*(0.0)*
OECD average	*499*	*(0.8)*	*103*	*(0.6)*	*501*	*(0.8)*	*97*	*(0.5)*	*–1.71*	*(0.8)*	*0.0*	*(0.0)*
Partner countries												
Brazil	374	(6.0)	106	(3.4)	368	(4.3)	95	(2.4)	5.21	(3.7)	0.1	(0.1)
Hong Kong-China	545	(6.2)	104	(3.8)	550	(4.0)	90	(2.6)	–5.06	(6.3)	0.1	(0.2)
Indonesia	358	(3.1)	72	(1.8)	365	(4.0)	74	(2.1)	**–7.30**	(3.0)	0.3	(0.2)
Latvia	481	(5.1)	97	(2.8)	484	(4.0)	87	(2.2)	–2.57	(4.6)	0.0	(0.1)
Liechtenstein	535	(6.6)	98	(5.8)	524	(5.9)	87	(5.2)	11.52	(9.8)	0.4	(0.7)
Macao-China	538	(4.2)	85	(4.2)	527	(3.2)	77	(2.9)	**11.22**	(5.5)	0.5	(0.5)
Russian Federation	480	(5.9)	104	(2.8)	477	(4.4)	93	(2.0)	2.30	(4.9)	0.0	(0.1)
Serbia	416	(3.8)	91	(2.0)	424	(3.9)	80	(2.2)	–7.39	(4.1)	0.2	(0.2)
Thailand	418	(3.9)	84	(2.3)	431	(3.1)	80	(2.0)	**–12.37**	(4.3)	0.6	(0.4)
Tunisia	346	(2.5)	80	(1.9)	343	(2.5)	79	(1.6)	2.71	(2.6)	0.0	(0.1)
Uruguay	412	(4.6)	116	(2.4)	409	(4.2)	108	(2.4)	2.73	(4.8)	0.0	(0.1)
United Kingdom[1]	m	m	m	m	m	m	m	m	m	m	m	m

	Problem solving								Reading		Mathematics	
	Percentage of males				Percentage of females				Gender difference in reading (M – F)		Gender difference in mathematics (M – F)	
	Mean score		Standard deviation		Mean score		Standard deviation		Gender difference		Gender difference	
	Below Level 1 (below 405 score points)	S.E.	At Level 3 (above 592 score points)	S.E.	Below Level 1 (below 405 score points)	S.E.	At Level 3 (above 592 score points)	S.E.	Score difference	S.E.	Score difference	S.E.
OECD countries												
Australia	10.8	(0.7)	25.7	(1.2)	7.5	(0.8)	25.7	(1.0)	**-39**	(3.6)	5	(3.8)
Austria	15.0	(1.2)	18.5	(1.5)	11.7	(1.3)	16.1	(1.4)	**-47**	(5.2)	8	(4.4)
Belgium	14.6	(1.1)	28.9	(1.3)	12.0	(1.1)	27.7	(1.1)	**-37**	(5.1)	8	(4.8)
Canada	9.2	(0.6)	27.9	(0.9)	6.8	(0.5)	24.1	(0.8)	**-32**	(2.0)	**11**	(2.1)
Czech Republic	11.7	(1.3)	23.5	(1.7)	12.0	(1.5)	19.5	(1.5)	**-31**	(4.9)	**15**	(5.1)
Denmark	10.2	(0.9)	21.0	(1.3)	10.4	(1.1)	19.4	(1.4)	**-25**	(2.9)	**17**	(3.2)
Finland	5.9	(0.8)	29.3	(1.2)	3.0	(0.4)	31.0	(1.2)	**-44**	(2.7)	**7**	(2.7)
France	13.0	(1.4)	24.0	(1.3)	10.1	(1.0)	21.7	(1.2)	**-38**	(4.5)	**9**	(4.2)
Germany	14.9	(1.3)	21.7	(1.6)	12.7	(1.2)	21.9	(1.8)	**-42**	(4.6)	**9**	(4.4)
Greece	33.3	(1.9)	8.4	(1.1)	31.4	(1.8)	5.8	(0.9)	**-37**	(4.1)	**19**	(3.6)
Hungary	16.9	(1.3)	17.2	(1.4)	14.7	(1.1)	17.3	(1.4)	**-31**	(3.8)	**8**	(3.5)
Iceland	17.1	(1.1)	12.2	(0.9)	6.8	(0.9)	17.9	(1.0)	**-58**	(3.5)	**-15**	(3.5)
Ireland	12.6	(1.2)	13.2	(0.9)	11.9	(1.3)	11.7	(1.2)	**-29**	(4.6)	**15**	(4.2)
Italy	26.6	(2.1)	12.1	(1.0)	22.5	(1.5)	9.2	(0.7)	**-39**	(6.0)	**18**	(5.9)
Japan	11.3	(1.3)	36.5	(2.3)	8.4	(1.2)	34.9	(1.5)	**-22**	(5.4)	8	(5.9)
Korea	5.2	(0.8)	34.6	(1.7)	4.7	(0.9)	29.5	(2.2)	**-21**	(5.6)	**23**	(6.8)
Luxembourg	17.7	(1.1)	16.1	(0.9)	15.8	(1.1)	12.5	(0.7)	**-33**	(3.4)	**17**	(2.8)
Mexico	57.2	(2.1)	1.6	(0.4)	58.2	(2.0)	1.0	(0.3)	**-21**	(4.4)	**11**	(3.9)
Netherlands	9.8	(1.2)	23.7	(1.4)	11.2	(1.5)	22.5	(1.4)	**-21**	(3.9)	5	(4.3)
New Zealand	10.9	(0.9)	28.8	(1.1)	8.6	(1.0)	27.9	(1.4)	**-28**	(4.4)	**14**	(3.9)
Norway	20.8	(1.1)	14.8	(1.1)	17.4	(1.2)	15.1	(1.2)	**-49**	(3.7)	6	(3.2)
Poland	18.8	(1.2)	13.0	(0.9)	15.6	(1.3)	10.5	(0.8)	**-40**	(3.7)	6	(3.1)
Portugal	25.6	(2.0)	10.1	(1.0)	21.8	(1.8)	7.3	(1.1)	**-36**	(3.3)	**12**	(3.3)
Slovak Republic	17.1	(1.6)	16.0	(1.2)	17.3	(1.6)	12.1	(1.0)	**-33**	(3.5)	**19**	(3.7)
Spain	21.7	(1.3)	12.2	(1.2)	17.8	(0.8)	11.2	(0.8)	**-39**	(3.9)	**9**	(3.0)
Sweden	12.8	(1.1)	16.2	(1.2)	10.8	(1.0)	18.7	(1.3)	**-37**	(3.2)	**7**	(3.3)
Switzerland	11.9	(1.1)	23.4	(2.0)	10.4	(0.9)	22.8	(1.4)	**-35**	(4.7)	**17**	(4.9)
Turkey	50.1	(2.9)	4.6	(1.5)	51.4	(2.8)	3.2	(1.1)	**-33**	(5.8)	**15**	(6.2)
United States	24.1	(1.3)	13.2	(1.1)	22.6	(1.3)	11.6	(0.9)	**-32**	(3.3)	**6**	(2.9)
OECD total	*22.2*	*(0.5)*	*18.1*	*(0.5)*	*20.5*	*(0.5)*	*16.5*	*(0.4)*	*-31*	*(1.4)*	*10*	*(1.4)*
OECD average	*18.0*	*(0.3)*	*19.0*	*(0.2)*	*16.0*	*(0.3)*	*17.6*	*(0.2)*	*-34*	*(0.8)*	*11*	*(0.8)*
Partner countries												
Brazil	62.6	(2.3)	2.4	(0.7)	64.7	(1.9)	0.9	(0.3)	**-35**	(3.9)	**16**	(4.1)
Hong Kong-China	9.8	(1.6)	36.1	(2.1)	5.9	(0.8)	33.9	(1.8)	**-32**	(5.5)	4	(6.6)
Indonesia	75.2	(1.8)	0.1	(0.1)	70.9	(2.2)	0.2	(0.1)	**-24**	(2.8)	3	(3.4)
Latvia	21.7	(1.9)	12.7	(1.3)	18.5	(1.7)	10.7	(1.1)	**-39**	(4.2)	3	(4.0)
Liechtenstein	9.8	(2.4)	31.2	(3.8)	10.0	(2.5)	22.7	(3.5)	-17	(11.9)	**29**	(10.9)
Macao-China	6.3	(1.3)	27.7	(2.7)	6.2	(1.2)	21.1	(2.3)	-13	(4.8)	**21**	(5.8)
Russian Federation	23.7	(2.2)	14.0	(1.5)	21.3	(1.6)	10.6	(1.1)	**-29**	(3.9)	**10**	(4.4)
Serbia	44.6	(2.0)	2.7	(0.6)	39.6	(2.0)	1.5	(0.3)	**-43**	(3.9)	1	(4.4)
Thailand	45.1	(2.1)	2.5	(0.5)	37.2	(1.7)	2.7	(0.7)	**-43**	(4.1)	**-4**	(4.2)
Tunisia	76.3	(1.3)	0.1	(0.1)	77.0	(1.3)	0.1	(0.1)	**-25**	(3.6)	**12**	(2.5)
Uruguay	46.2	(1.9)	5.5	(0.8)	47.4	(2.0)	4.0	(0.7)	**-39**	(4.7)	**12**	(4.2)
United Kingdom[1]	m	m	m	m	m	m	m	m	m	m	m	m

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.2
International socio-economic index of occupational status (HISEI) and performance on the problem-solving scale, by national quarters of the index
Results based on students' self-reports

		International socio-economic index of occupational status (highest of the father's or mother's)									
		All students		Bottom quarter		Second quarter		Third quarter		Top quarter	
		Mean index	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**
OECD countries	Australia	52.6	(0.30)	31.6	(0.14)	48.0	(0.07)	58.3	(0.11)	72.5	(0.14)
	Austria	47.1	(0.52)	27.3	(0.19)	40.9	(0.11)	51.4	(0.12)	68.7	(0.28)
	Belgium	50.6	(0.38)	29.0	(0.13)	44.5	(0.13)	56.4	(0.13)	72.4	(0.16)
	Canada	52.6	(0.27)	31.7	(0.11)	47.7	(0.08)	58.1	(0.09)	72.9	(0.15)
	Czech Republic	50.1	(0.34)	32.3	(0.18)	45.7	(0.12)	52.5	(0.05)	69.7	(0.23)
	Denmark	49.3	(0.45)	29.4	(0.19)	44.2	(0.11)	53.2	(0.07)	70.3	(0.29)
	Finland	50.2	(0.36)	28.7	(0.12)	43.4	(0.16)	56.4	(0.14)	72.4	(0.18)
	France	48.7	(0.47)	27.6	(0.20)	42.3	(0.15)	53.6	(0.05)	71.2	(0.26)
	Germany	49.3	(0.42)	29.5	(0.17)	42.6	(0.14)	53.7	(0.06)	71.5	(0.25)
	Greece	46.9	(0.72)	26.9	(0.13)	38.8	(0.13)	51.8	(0.07)	70.3	(0.39)
	Hungary	48.6	(0.33)	30.2	(0.18)	42.3	(0.08)	51.6	(0.11)	70.2	(0.20)
	Iceland	53.7	(0.26)	31.5	(0.20)	48.0	(0.13)	61.7	(0.19)	73.7	(0.25)
	Ireland	48.3	(0.49)	28.5	(0.17)	42.2	(0.11)	52.7	(0.08)	70.0	(0.29)
	Italy	46.8	(0.38)	26.9	(0.16)	40.3	(0.11)	50.6	(0.05)	69.5	(0.38)
	Japan	50.0	(0.31)	33.4	(0.17)	43.9	(0.04)	50.6	(0.08)	72.0	(0.25)
	Korea	46.3	(0.36)	28.9	(0.20)	43.5	(0.09)	49.4	(0.06)	63.5	(0.43)
	Luxembourg	48.2	(0.22)	27.3	(0.15)	42.1	(0.13)	52.8	(0.06)	70.5	(0.24)
	Mexico	40.1	(0.68)	22.2	(0.12)	28.9	(0.04)	42.1	(0.28)	67.3	(0.25)
	Netherlands	51.3	(0.38)	30.9	(0.26)	45.4	(0.15)	56.9	(0.20)	71.8	(0.25)
	New Zealand	51.5	(0.36)	30.1	(0.19)	46.2	(0.12)	56.8	(0.17)	72.7	(0.26)
	Norway	54.6	(0.39)	35.0	(0.20)	49.0	(0.12)	60.6	(0.16)	73.9	(0.21)
	Poland	45.0	(0.34)	26.9	(0.21)	39.5	(0.11)	49.1	(0.10)	64.4	(0.34)
	Portugal	43.1	(0.54)	26.4	(0.14)	33.9	(0.08)	46.6	(0.19)	65.5	(0.53)
	Slovak Republic	48.8	(0.40)	29.3	(0.17)	41.4	(0.09)	53.1	(0.10)	71.5	(0.21)
	Spain	44.3	(0.58)	26.2	(0.13)	35.5	(0.14)	49.3	(0.11)	66.1	(0.39)
	Sweden	50.6	(0.38)	30.4	(0.18)	44.1	(0.14)	56.1	(0.17)	71.9	(0.21)
	Switzerland	49.3	(0.43)	29.4	(0.14)	43.1	(0.14)	53.5	(0.08)	71.1	(0.27)
	Turkey	41.6	(0.75)	23.7	(0.29)	33.6	(0.15)	47.2	(0.10)	61.8	(0.77)
	United States	54.6	(0.37)	32.6	(0.21)	49.9	(0.15)	61.4	(0.12)	74.3	(0.21)
	OECD total	*49.2*	*(0.15)*	*28.1*	*(0.07)*	*42.5*	*(0.07)*	*54.1*	*(0.08)*	*71.9*	*(0.11)*
	OECD average	*48.8*	*(0.08)*	*28.2*	*(0.04)*	*42.3*	*(0.08)*	*53.2*	*(0.09)*	*71.2*	*(0.13)*
Partner countries	Brazil	40.1	(0.64)	21.7	(0.31)	32.4	(0.09)	44.4	(0.17)	62.1	(0.60)
	Hong Kong-China	41.1	(0.45)	25.9	(0.14)	34.9	(0.07)	45.1	(0.13)	58.7	(0.37)
	Indonesia	33.6	(0.61)	16.0	(0.00)	24.1	(0.15)	34.6	(0.33)	59.9	(0.42)
	Latvia	50.3	(0.52)	29.1	(0.23)	44.2	(0.16)	54.8	(0.14)	73.0	(0.30)
	Liechtenstein	50.7	(0.75)	30.8	(0.63)	47.4	(0.52)	55.0	(0.09)	70.0	(0.67)
	Macao-China	39.4	(0.40)	25.8	(0.32)	34.4	(0.12)	41.7	(0.25)	55.9	(0.52)
	Russian Federation	49.9	(0.38)	30.8	(0.16)	40.9	(0.10)	54.2	(0.21)	73.6	(0.20)
	Serbia	48.1	(0.53)	28.3	(0.20)	41.2	(0.12)	51.4	(0.11)	71.4	(0.38)
	Thailand	36.0	(0.43)	22.1	(0.14)	26.7	(0.13)	35.6	(0.13)	59.6	(0.41)
	Tunisia	37.5	(0.60)	18.0	(0.17)	29.2	(0.18)	39.6	(0.19)	63.1	(0.44)
	Uruguay	46.1	(0.48)	25.2	(0.16)	37.8	(0.15)	50.8	(0.12)	70.8	(0.36)
	United Kingdom[1]	49.6	(0.39)	28.5	(0.14)	43.0	(0.14)	55.5	(0.11)	71.6	(0.19)

		Performance on the problem-solving scale, by national quarters of the international socio-economic index of occupational status								Change in the problem-solving score per 16.4 units of the international socio-economic index of occupational status		Explained variance in student performance (r-squared × 100)	
		Bottom quarter		Second quarter		Third quarter		Top quarter					
		Mean score	**S.E.**	**Mean score**	**S.E.**	**Mean score**	**S.E.**	**Mean score**	**S.E.**	**Effect**	**S.E.**	**%**	**S.E.**
OECD countries	Australia	**498**	(2.5)	526	(2.6)	543	(2.3)	**566**	(2.8)	**26.6**	(1.3)	8.3	(0.7)
	Austria	**468**	(4.0)	493	(3.5)	522	(2.8)	**549**	(4.2)	**30.2**	(1.8)	11.0	(1.2)
	Belgium	**481**	(3.7)	524	(3.3)	550	(2.8)	**580**	(3.1)	**36.0**	(1.7)	13.9	(1.2)
	Canada	**505**	(2.4)	528	(2.2)	539	(2.0)	**565**	(2.4)	**23.5**	(1.2)	6.8	(0.7)
	Czech Republic	**490**	(4.1)	504	(3.6)	530	(3.3)	**566**	(4.3)	**33.3**	(2.0)	10.8	(1.2)
	Denmark	**488**	(3.4)	507	(3.3)	527	(3.7)	**552**	(3.2)	**25.2**	(1.7)	7.6	(1.0)
	Finland	**521**	(2.7)	542	(2.6)	555	(2.8)	**577**	(3.1)	**19.6**	(1.4)	6.1	(0.8)
	France	**476**	(4.1)	511	(4.4)	534	(3.1)	**567**	(3.8)	**32.8**	(2.1)	13.6	(1.5)
	Germany	**473**	(4.6)	514	(3.4)	536	(4.1)	**567**	(3.7)	**34.8**	(1.8)	14.3	(1.2)
	Greece	**411**	(4.4)	441	(4.4)	454	(4.9)	**498**	(5.0)	**30.1**	(2.2)	10.0	(1.5)
	Hungary	**457**	(4.1)	482	(3.6)	522	(3.4)	**558**	(4.3)	**41.4**	(2.3)	17.1	(1.6)
	Iceland	**489**	(3.0)	500	(3.0)	508	(2.5)	**529**	(3.3)	**13.7**	(1.5)	2.8	(0.6)
	Ireland	**466**	(3.5)	495	(3.1)	509	(3.1)	**536**	(3.4)	**27.2**	(1.9)	11.3	(1.5)
	Italy	**433**	(4.7)	461	(4.2)	485	(3.6)	**507**	(4.3)	**28.3**	(2.1)	8.1	(1.0)
	Japan	**520**	(5.1)	551	(5.0)	556	(4.7)	**580**	(6.0)	**21.5**	(3.2)	3.6	(0.9)
	Korea	**526**	(4.2)	555	(3.5)	555	(3.5)	**573**	(5.4)	**22.3**	(2.9)	4.5	(1.1)
	Luxembourg	**450**	(2.9)	484	(3.2)	511	(2.6)	**539**	(3.2)	**32.4**	(1.6)	12.9	(1.2)
	Mexico	**352**	(5.3)	371	(4.5)	392	(4.7)	**433**	(5.7)	**28.6**	(2.3)	11.1	(1.7)
	Netherlands	**482**	(4.5)	518	(4.1)	540	(3.7)	**564**	(3.8)	**31.9**	(2.1)	12.4	(1.4)
	New Zealand	**492**	(3.8)	525	(3.4)	543	(3.5)	**573**	(3.6)	**30.2**	(1.8)	10.1	(1.2)
	Norway	**449**	(3.8)	483	(3.4)	502	(4.2)	**529**	(3.9)	**30.9**	(1.8)	8.8	(0.9)
	Poland	**455**	(4.0)	475	(3.3)	494	(3.5)	**528**	(3.5)	**33.5**	(1.9)	11.6	(1.2)
	Portugal	**433**	(5.7)	454	(4.6)	486	(3.9)	**515**	(4.2)	**34.7**	(2.1)	13.7	(1.5)
	Slovak Republic	**450**	(4.6)	478	(3.4)	516	(3.6)	**541**	(3.7)	**34.2**	(1.9)	14.1	(1.2)
	Spain	**449**	(3.8)	471	(3.0)	495	(3.6)	**521**	(3.7)	**27.9**	(1.6)	8.9	(1.0)
	Sweden	**477**	(3.5)	500	(2.8)	517	(3.2)	**545**	(3.5)	**26.7**	(1.6)	9.2	(1.0)
	Switzerland	**480**	(3.6)	520	(3.6)	532	(4.7)	**561**	(3.4)	**28.4**	(1.6)	9.1	(0.9)
	Turkey	**384**	(5.3)	395	(6.1)	405	(6.6)	**461**	(11.2)	**34.1**	(5.1)	11.2	(2.9)
	United States	**442**	(3.6)	473	(3.8)	492	(3.8)	**526**	(3.6)	**29.9**	(1.5)	9.6	(0.9)
	OECD total	***455***	*(1.5)*	*484*	*(1.5)*	*502*	*(1.6)*	***533***	*(1.3)*	***33.5***	*(0.8)*	*13.3*	*(0.7)*
	OECD average	***465***	*(4.0)*	*492*	*(3.6)*	*512*	*(3.6)*	***542***	*(4.1)*	***32.9***	*(0.4)*	*11.2*	*(0.2)*
Partner countries	Brazil	**333**	(5.4)	367	(5.6)	386	(4.9)	**422**	(7.7)	**36.6**	(3.4)	13.2	(2.2)
	Hong Kong-China	**526**	(5.6)	546	(4.6)	559	(3.5)	**573**	(4.9)	**24.4**	(2.6)	4.4	(0.9)
	Indonesia	**336**	(3.4)	355	(3.6)	362	(4.5)	**399**	(4.8)	**22.2**	(2.0)	10.2	(1.6)
	Latvia	**459**	(4.4)	472	(5.1)	492	(4.6)	**511**	(5.5)	**19.6**	(2.1)	4.7	(1.1)
	Liechtenstein	**479**	(9.1)	527	(10.1)	538	(10.2)	**582**	(9.0)	**38.6**	(5.1)	14.6	(3.7)
	Macao-China	**526**	(4.9)	527	(5.8)	535	(6.5)	**545**	(6.7)	**12.1**	(3.8)	1.2	(0.8)
	Russian Federation	**450**	(5.7)	468	(5.3)	485	(5.1)	**514**	(5.0)	**23.5**	(2.1)	6.0	(1.0)
	Serbia	**390**	(3.9)	411	(3.8)	432	(3.8)	**456**	(4.6)	**24.7**	(2.0)	8.6	(1.3)
	Thailand	**404**	(3.2)	409	(3.4)	431	(3.5)	**464**	(4.7)	**25.9**	(2.3)	9.0	(1.5)
	Tunisia	**316**	(3.4)	327	(2.8)	356	(3.9)	**388**	(4.8)	**26.9**	(2.0)	13.4	(2.0)
	Uruguay	**374**	(5.3)	404	(5.0)	418	(4.7)	**474**	(4.7)	**34.9**	(2.2)	11.7	(1.3)
	United Kingdom[1]	**473**	(3.0)	504	(3.3)	520	(3.8)	**553**	(3.6)	**29.4**	(1.5)	10.5	(1.1)

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.3
Index of highest educational level of parents (HISCED)[1] and performance on the problem-solving scale, by national quarters of the index
Results based on students' self-reports

	Index of highest educational level of parents (HISCED)									
	All students		Bottom quarter		Second quarter		Third quarter		Top quarter	
	Mean index	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**	**Mean index**	**S.E.**
OECD countries										
Australia	4.62	(0.02)	2.71	(0.03)	4.17	(0.01)	5.59	(0.01)	max	
Austria	4.07	(0.03)	2.68	(0.03)	3.33	(0.02)	4.73	(0.01)	5.55	(0.02)
Belgium	4.64	(0.03)	2.88	(0.04)	4.29	(0.01)	5.41	(0.01)	max	
Canada	4.93	(0.02)	3.62	(0.02)	4.54	(0.01)	5.56	(0.01)	max	
Czech Republic	4.23	(0.03)	2.91	(0.01)	3.93	(0.01)	4.07	(0.01)	max	
Denmark	4.47	(0.04)	2.69	(0.05)	4.34	(0.02)	5.00	(0.00)	5.85	(0.01)
Finland	4.78	(0.02)	3.06	(0.03)	4.67	(0.01)	5.40	(0.02)	max	
France	3.98	(0.04)	2.00	(0.03)	3.36	(0.02)	4.54	(0.02)	max	
Germany	4.02	(0.04)	1.90	(0.04)	3.53	(0.02)	4.65	(0.02)	5.99	(0.00)
Greece	4.16	(0.06)	1.84	(0.03)	4.00	(0.00)	4.80	(0.02)	max	
Hungary	4.24	(0.03)	2.70	(0.02)	3.94	(0.01)	4.33	(0.01)	max	
Iceland	4.29	(0.02)	2.49	(0.03)	4.00	(0.00)	4.67	(0.02)	max	
Ireland	4.22	(0.04)	2.48	(0.06)	4.00	(0.00)	4.60	(0.02)	5.79	(0.01)
Italy	3.86	(0.03)	1.83	(0.01)	3.43	(0.02)	4.38	(0.01)	5.81	(0.01)
Japan	4.78	(0.03)	3.15	(0.06)	4.32	(0.01)	5.65	(0.02)	max	
Korea	4.07	(0.04)	1.82	(0.03)	3.74	(0.01)	4.71	(0.03)	max	
Luxembourg	4.09	(0.03)	1.42	(0.04)	4.01	(0.01)	5.00	(0.00)	5.94	(0.01)
Mexico	2.91	(0.07)	0.50	(0.02)	1.71	(0.02)	3.71	(0.05)	5.70	(0.02)
Netherlands	4.55	(0.04)	2.38	(0.06)	4.00	(0.00)	5.80	(0.02)	max	
New Zealand	4.24	(0.03)	2.26	(0.04)	4.00	(0.00)	4.79	(0.01)	5.91	(0.01)
Norway	4.75	(0.02)	3.48	(0.03)	4.51	(0.02)	5.01	(0.00)	max	
Poland	4.10	(0.02)	2.90	(0.02)	3.99	(0.00)	4.01	(0.00)	5.50	(0.03)
Portugal	2.74	(0.06)	0.27	(0.01)	1.46	(0.02)	3.48	(0.03)	5.74	(0.02)
Slovak Republic	4.26	(0.03)	3.05	(0.04)	4.00	(0.00)	4.06	(0.01)	5.93	(0.01)
Spain	3.66	(0.07)	0.87	(0.01)	3.21	(0.03)	4.57	(0.02)	max	
Sweden	4.66	(0.03)	2.69	(0.04)	4.43	(0.01)	5.52	(0.02)	max	
Switzerland	3.88	(0.04)	1.87	(0.02)	3.14	(0.01)	4.66	(0.01)	5.86	(0.01)
Turkey	2.81	(0.09)	0.83	(0.02)	1.59	(0.02)	3.49	(0.03)	5.35	(0.04)
United States	4.69	(0.03)	3.41	(0.06)	4.01	(0.00)	5.36	(0.02)	max	
OECD total	*4.18*	*(0.01)*	*1.88*	*(0.01)*	*3.99*	*(0.00)*	*4.86*	*(0.01)*	*max*	
OECD average	*4.16*	*(0.01)*	*1.92*	*(0.01)*	*3.92*	*(0.00)*	*4.78*	*(0.02)*	*max*	
Partner countries										
Brazil	3.66	(0.07)	0.87	(0.03)	2.62	(0.04)	5.16	(0.03)	max	
Hong Kong-China	2.59	(0.04)	0.94	(0.01)	2.00	(0.00)	2.79	(0.01)	4.62	(0.04)
Indonesia	2.83	(0.06)	0.72	(0.02)	1.58	(0.02)	3.68	(0.02)	5.33	(0.03)
Latvia	4.83	(0.03)	3.73	(0.02)	4.39	(0.02)	5.18	(0.01)	max	
Liechtenstein	3.92	(0.07)	2.02	(0.08)	3.12	(0.03)	4.75	(0.04)	5.83	(0.04)
Macao-China	2.58	(0.05)	0.65	(0.04)	1.93	(0.02)	3.27	(0.07)	4.46	(0.05)
Russian Federation	4.83	(0.03)	3.90	(0.01)	4.00	(0.00)	5.42	(0.03)	max	
Serbia	4.19	(0.04)	2.68	(0.02)	3.27	(0.02)	5.00	(0.00)	5.83	(0.02)
Thailand	2.39	(0.05)	0.93	(0.02)	1.01	(0.00)	2.58	(0.03)	5.06	(0.03)
Tunisia	2.46	(0.06)	0.54	(0.02)	1.43	(0.02)	2.97	(0.03)	4.90	(0.03)
Uruguay	3.88	(0.05)	1.31	(0.02)	3.12	(0.03)	5.08	(0.01)	max	
United Kingdom[1]	4.20	(0.03)	2.53	(0.03)	3.71	(0.01)	4.68	(0.01)	5.89	(0.01)

	Performance on the problem-solving scale, by national quarters of the index of highest educational level of parents (HISCED)								Change in the problem-solving score per unit of the index of highest educational level of parents		Explained variance in student performance (r-squared × 100)	
	Bottom quarter		Second quarter		Third quarter		Top quarter					
	Mean score	**S.E.**	**Mean score**	**S.E.**	**Mean score**	**S.E.**	**Mean score**	**S.E.**	**Effect**	**S.E.**	**%**	**S.E.**
OECD countries												
Australia	**508**	(3.1)	516	(3.4)	543	(2.8)	**559**	(3.0)	**14.3**	(1.1)	5.2	(0.7)
Austria	**485**	(4.5)	507	(4.4)	510	(4.2)	**531**	(4.7)	**15.2**	(1.6)	4.5	(1.0)
Belgium	**495**	(4.3)	528	(3.2)	552	(3.1)	**561**	(3.2)	**21.5**	(1.2)	9.1	(1.0)
Canada	**512**	(2.0)	526	(2.2)	545	(3.0)	**554**	(2.8)	**17.1**	(0.9)	4.3	(0.4)
Czech Republic	**481**	(3.6)	522	(3.8)	525	(3.6)	**565**	(4.5)	**26.1**	(1.7)	11.3	(1.3)
Denmark	**487**	(3.6)	515	(3.6)	534	(3.7)	**542**	(3.8)	**18.6**	(1.5)	7.9	(1.2)
Finland	**530**	(2.7)	544	(2.9)	554	(2.6)	**566**	(2.8)	**10.6**	(1.0)	2.8	(0.5)
France	**487**	(4.8)	519	(3.8)	536	(4.2)	**551**	(4.0)	**16.0**	(1.3)	7.8	(1.2)
Germany	**475**	(4.7)	523	(4.4)	528	(5.0)	**565**	(4.2)	**20.6**	(1.2)	13.6	(1.5)
Greece	**416**	(4.5)	448	(4.2)	448	(5.5)	**482**	(6.3)	**14.5**	(1.6)	5.5	(1.1)
Hungary	**449**	(4.3)	495	(3.5)	501	(3.5)	**560**	(4.4)	**33.1**	(1.7)	18.9	(1.6)
Iceland	**487**	(3.5)	502	(3.0)	505	(3.7)	**529**	(2.9)	**11.5**	(1.0)	3.5	(0.6)
Ireland	**473**	(3.3)	494	(3.3)	503	(3.3)	**526**	(3.5)	**15.8**	(1.1)	7.6	(1.1)
Italy	**431**	(4.9)	474	(3.8)	478	(4.0)	**497**	(3.9)	**15.7**	(1.4)	5.7	(0.9)
Japan	**519**	(6.1)	529	(5.9)	566	(5.5)	**576**	(5.2)	**19.4**	(2.7)	5.7	(1.3)
Korea	**523**	(4.5)	544	(3.4)	555	(3.6)	**582**	(5.4)	**13.5**	(1.4)	6.6	(1.2)
Luxembourg	**466**	(3.1)	494	(3.4)	500	(3.2)	**533**	(3.1)	**13.9**	(0.8)	7.3	(0.8)
Mexico	**337**	(4.6)	372	(4.4)	418	(5.5)	**413**	(5.7)	**15.0**	(1.3)	10.6	(1.6)
Netherlands	**502**	(5.5)	516	(3.8)	543	(4.5)	**545**	(6.0)	**12.6**	(1.4)	5.5	(1.1)
New Zealand	**513**	(3.7)	537	(3.1)	538	(3.3)	**568**	(4.0)	**15.4**	(1.2)	5.7	(0.9)
Norway	**469**	(4.3)	490	(4.5)	506	(4.3)	**506**	(4.3)	**15.7**	(1.5)	2.8	(0.5)
Poland	**453**	(4.2)	485	(4.1)	485	(3.8)	**524**	(3.5)	**26.3**	(1.9)	9.1	(1.2)
Portugal	**440**	(4.7)	463	(4.5)	486	(3.9)	**496**	(6.0)	**10.5**	(1.0)	6.2	(1.1)
Slovak Republic	**453**	(6.0)	489	(3.9)	488	(3.7)	**540**	(3.4)	**29.9**	(2.3)	13.3	(1.8)
Spain	**449**	(4.6)	481	(3.0)	484	(3.9)	**522**	(3.8)	**12.4**	(1.0)	7.1	(1.1)
Sweden	**490**	(3.4)	516	(3.3)	520	(4.4)	**520**	(3.6)	**11.1**	(1.1)	3.3	(0.6)
Switzerland	**480**	(3.3)	537	(3.1)	526	(3.4)	**550**	(5.1)	**15.3**	(1.1)	6.8	(0.9)
Turkey	**373**	(4.6)	388	(5.5)	415	(6.3)	**456**	(11.7)	**19.6**	(2.5)	14.5	(3.0)
United States	**453**	(4.2)	466	(4.6)	486	(4.3)	**511**	(4.5)	**21.3**	(1.2)	7.1	(0.9)
OECD total	*443*	*(1.8)*	*489*	*(1.4)*	*501*	*(2.0)*	*531*	*(1.5)*	*22.7*	*(0.5)*	*12.0*	*(0.5)*
OECD average	*459*	*(1.2)*	*500*	*(0.7)*	*510*	*(0.9)*	*540*	*(0.8)*	*20.2*	*(0.3)*	*10.4*	*(0.3)*
Partner countries												
Brazil	**339**	(5.0)	375	(5.0)	391	(6.4)	**385**	(7.1)	**9.5**	(1.3)	4.1	(1.0)
Hong Kong-China	**528**	(4.8)	542	(6.1)	562	(4.9)	**564**	(5.2)	**9.4**	(1.5)	2.0	(0.6)
Indonesia	**346**	(3.1)	352	(3.8)	374	(3.8)	**375**	(6.0)	**6.5**	(1.2)	2.8	(1.0)
Latvia	**469**	(5.5)	474	(5.8)	481	(5.7)	**510**	(5.7)	**15.9**	(2.6)	2.7	(0.9)
Liechtenstein	**501**	(9.3)	520	(12.8)	541	(9.0)	**563**	(11.9)	**16.2**	(3.5)	7.3	(3.4)
Macao-China	**525**	(6.5)	526	(7.2)	538	(6.7)	**540**	(4.8)	3.6	(2.0)	0.5	(0.6)
Russian Federation	**460**	(5.7)	460	(5.2)	491	(6.7)	**505**	(5.8)	**21.6**	(2.0)	5.1	(0.9)
Serbia	**404**	(4.3)	421	(4.6)	409	(4.4)	**448**	(5.1)	**11.1**	(1.5)	3.0	(0.8)
Thailand	**410**	(3.3)	409	(3.7)	424	(3.8)	**458**	(5.1)	**12.1**	(1.3)	7.2	(1.4)
Tunisia	**323**	(3.0)	335	(2.9)	350	(3.0)	**372**	(5.2)	**10.7**	(1.3)	5.8	(1.4)
Uruguay	**370**	(5.7)	400	(5.2)	417	(5.1)	**458**	(5.6)	**16.9**	(1.4)	8.4	(1.3)
United Kingdom[1]	**489**	(4.0)	502	(3.3)	510	(4.1)	**549**	(3.6)	**17.8**	(1.4)	6.7	(1.0)

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Highest Educational level of Parents (HISCED) corresponds to the higher level of education (ISCED) of either parent.
2. Response rate too low to ensure comparability (see Annex A3).

Table 5.4

Index of possessions related to "classical" culture in the family home and performance on the problem-solving scale, by national quarters of the index

Results based on students' self-reports

		Index of possessions related to "classical" culture in the family home									
		All students		Bottom quarter		Second quarter		Third quarter		Top quarter	
		Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.
OECD countries	Australia	–0.12	(0.01)	min		–0.64	(0.01)	0.13	(0.01)	1.31	(0.00)
	Austria	–0.05	(0.03)	min		–0.48	(0.01)	0.28	(0.01)	1.29	(0.01)
	Belgium	–0.30	(0.02)	min		–0.94	(0.01)	–0.05	(0.01)	1.08	(0.01)
	Canada	0.00	(0.01)	min		–0.40	(0.01)	0.32	(0.01)	max	
	Czech Republic	0.26	(0.02)	–1.00	(0.02)	–0.02	(0.01)	0.71	(0.01)	max	
	Denmark	–0.01	(0.03)	min		–0.45	(0.01)	0.35	(0.00)	max	
	Finland	0.11	(0.02)	min		–0.28	(0.01)	0.65	(0.01)	max	
	France	–0.05	(0.02)	min		–0.44	(0.01)	0.30	(0.01)	1.22	(0.01)
	Germany	0.00	(0.02)	min		–0.44	(0.01)	0.37	(0.01)	max	
	Greece	0.23	(0.03)	–0.94	(0.01)	–0.07	(0.01)	0.59	(0.01)	max	
	Hungary	0.31	(0.02)	–0.97	(0.02)	0.16	(0.01)	0.69	(0.01)	max	
	Iceland	0.79	(0.01)	–0.42	(0.02)	0.90	(0.02)	1.35	(0.00)	max	
	Ireland	–0.26	(0.02)	min		–0.85	(0.02)	0.01	(0.01)	1.07	(0.01)
	Italy	0.19	(0.02)	–1.18	(0.01)	–0.08	(0.01)	0.67	(0.01)	max	
	Japan	–0.43	(0.02)	min		–1.12	(0.01)	–0.18	(0.01)	0.85	(0.01)
	Korea	0.16	(0.02)	–1.14	(0.01)	–0.11	(0.01)	0.55	(0.01)	max	
	Luxembourg	–0.03	(0.01)	min		–0.51	(0.01)	0.31	(0.01)	max	
	Mexico	–0.68	(0.03)	min		–1.28	(0.00)	–0.65	(0.02)	0.49	(0.02)
	Netherlands	–0.31	(0.02)	min		–0.78	(0.02)	–0.16	(0.01)	0.96	(0.02)
	New Zealand	–0.18	(0.02)	min		–0.62	(0.01)	0.06	(0.01)	1.11	(0.01)
	Norway	0.15	(0.02)	min		–0.30	(0.01)	0.84	(0.02)	max	
	Poland	0.25	(0.02)	–0.84	(0.02)	–0.04	(0.01)	0.53	(0.01)	max	
	Portugal	–0.08	(0.03)	min		–0.55	(0.01)	0.27	(0.01)	1.24	(0.01)
	Slovak Republic	0.35	(0.02)	–0.93	(0.02)	0.10	(0.01)	0.88	(0.01)	max	
	Spain	0.15	(0.02)	–1.17	(0.01)	–0.11	(0.01)	0.54	(0.01)	max	
	Sweden	0.10	(0.02)	–1.26	(0.00)	–0.28	(0.00)	0.59	(0.01)	max	
	Switzerland	–0.37	(0.03)	min		–1.02	(0.01)	–0.13	(0.01)	0.95	(0.02)
	Turkey	–0.11	(0.03)	min		–0.51	(0.02)	0.22	(0.01)	1.12	(0.01)
	United States	–0.04	(0.02)	min		–0.57	(0.01)	0.34	(0.01)	max	
	OECD total	*–0.10*	*(0.01)*	*–1.28*	*(0.00)*	*–0.62*	*(0.01)*	*0.21*	*(0.01)*	*1.29*	*(0.01)*
	OECD average	*0.00*	*(0.00)*	*–1.28*	*(0.00)*	*–0.45*	*(0.00)*	*0.38*	*(0.02)*	*1.35*	*(0.00)*
Partner countries	Brazil	–0.33	(0.02)	–1.28	(0.00)	–0.83	(0.02)	–0.06	(0.01)	0.86	(0.02)
	Hong Kong-China	–0.44	(0.03)	min		–1.04	(0.01)	–0.22	(0.01)	0.78	(0.02)
	Indonesia	–0.65	(0.02)	min		min		–0.51	(0.01)	0.46	(0.02)
	Latvia	0.40	(0.02)	–0.91	(0.02)	0.25	(0.01)	0.92	(0.02)	max	
	Liechtenstein	–0.27	(0.05)	min		–0.85	(0.05)	–0.04	(0.03)	1.09	(0.05)
	Macao-China	–0.50	(0.02)	min		–1.16	(0.02)	–0.24	(0.01)	0.69	(0.03)
	Russian Federation	0.48	(0.02)	–0.67	(0.02)	0.38	(0.00)	0.85	(0.01)	max	
	Serbia	0.14	(0.03)	min		–0.22	(0.01)	0.73	(0.01)	max	
	Thailand	–0.21	(0.02)	min		–0.62	(0.01)	0.05	(0.01)	1.02	(0.01)
	Tunisia	–0.63	(0.02)	min		min		–0.47	(0.01)	0.50	(0.02)
	Uruguay	0.07	(0.02)	–1.21	(0.01)	–0.22	(0.01)	0.38	(0.00)	1.32	(0.00)
	United Kingdom[1]	–0.03	(0.02)	min		–0.61	(0.01)	0.40	(0.01)	max	

		Performance on the problem-solving scale, by national quarters of the index of possessions related to "classical" culture in the family home								Change in the problem-solving score per unit of this index		Explained variance in student performance (r-squared × 100)	
		Bottom quarter		Second quarter		Third quarter		Top quarter					
		Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Mean score	S.E.	Effect	S.E.	%	S.E.
OECD countries	Australia	**508**	(3.0)	519	(3.0)	532	(2.4)	**561**	(2.6)	**20.1**	(1.2)	4.9	(0.5)
	Austria	**477**	(4.0)	491	(3.9)	513	(4.2)	**551**	(4.1)	**29.1**	(1.7)	10.3	(1.1)
	Belgium	**494**	(3.9)	506	(3.6)	541	(3.2)	**571**	(2.8)	**32.6**	(1.7)	9.7	(0.8)
	Canada	**510**	(2.4)	529	(2.5)	538	(2.5)	**557**	(2.6)	**17.5**	(1.3)	3.9	(0.5)
	Czech Republic	**491**	(4.4)	518	(4.3)	534	(3.6)	**547**	(4.0)	**23.8**	(1.6)	6.4	(0.8)
	Denmark	**474**	(3.2)	502	(3.2)	535	(3.2)	**557**	(3.2)	**32.6**	(1.5)	13.7	(1.1)
	Finland	**521**	(3.1)	542	(2.8)	561	(3.5)	**568**	(3.3)	**18.2**	(1.4)	5.1	(0.7)
	France	**477**	(4.8)	512	(4.3)	539	(3.8)	**553**	(3.6)	**30.8**	(2.2)	10.3	(1.2)
	Germany	**495**	(4.1)	505	(4.5)	514	(4.4)	**561**	(3.9)	**23.9**	(1.5)	6.6	(0.8)
	Greece	**412**	(4.7)	445	(4.5)	462	(5.1)	**476**	(5.8)	**27.8**	(2.6)	6.6	(1.2)
	Hungary	**441**	(3.9)	501	(4.1)	525	(4.5)	**537**	(4.0)	**42.0**	(2.1)	16.8	(1.4)
	Iceland	**482**	(3.7)	507	(3.6)	516	(3.0)	**515**	(3.2)	**20.1**	(2.0)	3.8	(0.7)
	Ireland	**481**	(4.2)	488	(4.6)	500	(3.6)	**524**	(3.4)	**18.6**	(1.5)	5.1	(0.8)
	Italy	**434**	(4.4)	466	(4.5)	482	(4.2)	**498**	(3.9)	**25.2**	(1.9)	6.0	(0.8)
	Japan	**521**	(5.3)	527	(5.5)	566	(5.1)	**576**	(5.4)	**25.9**	(2.8)	5.1	(1.0)
	Korea	**521**	(3.9)	544	(4.3)	560	(3.6)	**578**	(4.8)	**23.1**	(2.4)	6.6	(1.2)
	Luxembourg	**474**	(3.0)	481	(3.3)	492	(2.8)	**528**	(2.8)	**20.5**	(1.4)	5.1	(0.6)
	Mexico	**364**	(4.7)	365	(4.5)	381	(4.9)	**427**	(6.3)	**35.7**	(3.1)	9.1	(1.6)
	Netherlands	**498**	(5.4)	513	(4.8)	524	(4.0)	**557**	(4.0)	**25.7**	(2.4)	7.0	(1.2)
	New Zealand	**507**	(3.8)	521	(4.1)	534	(5.0)	**572**	(3.6)	**26.8**	(1.7)	7.1	(0.9)
	Norway	**446**	(3.8)	480	(3.6)	511	(3.9)	**525**	(4.6)	**29.5**	(1.5)	10.0	(1.0)
	Poland	**455**	(4.1)	483	(3.5)	501	(4.5)	**508**	(4.2)	**25.5**	(2.2)	6.0	(0.9)
	Portugal	**432**	(5.6)	454	(4.4)	485	(4.4)	**509**	(4.8)	**30.7**	(2.1)	10.6	(1.3)
	Slovak Republic	**459**	(5.7)	485	(4.3)	506	(3.5)	**516**	(3.4)	**25.0**	(2.4)	6.4	(1.1)
	Spain	**447**	(4.0)	476	(3.7)	496	(3.6)	**511**	(3.9)	**25.7**	(1.8)	7.0	(0.9)
	Sweden	**470**	(3.5)	497	(3.1)	528	(3.7)	**541**	(3.6)	**27.9**	(1.9)	10.0	(1.2)
	Switzerland	**511**	(4.0)	511	(4.4)	516	(4.0)	**547**	(4.9)	**17.2**	(1.7)	3.0	(0.6)
	Turkey	**377**	(5.2)	390	(5.7)	414	(6.7)	**449**	(9.5)	**30.7**	(3.9)	9.0	(1.7)
	United States	**441**	(3.9)	458	(4.2)	489	(3.8)	**523**	(4.2)	**31.4**	(1.6)	10.9	(1.0)
	OECD total	*453*	*(2.0)*	*475*	*(1.6)*	*502*	*(1.2)*	*530*	*(1.5)*	*30.3*	*(0.7)*	*8.1*	*(0.3)*
	OECD average	*468*	*(1.1)*	*489*	*(0.8)*	*511*	*(0.8)*	*534*	*(0.8)*	*25.3*	*(0.4)*	*6.4*	*(0.2)*
Partner countries	Brazil	**358**	(5.3)	361	(6.0)	372	(6.7)	**397**	(6.7)	**19.1**	(2.3)	2.9	(0.7)
	Hong Kong-China	**526**	(5.8)	534	(5.9)	556	(5.5)	**575**	(5.2)	**22.9**	(2.6)	4.2	(0.9)
	Indonesia	**358**	(4.4)	357	(4.4)	364	(4.6)	**367**	(4.3)	**5.9**	(1.7)	0.4	(0.2)
	Latvia	**443**	(5.7)	486	(4.8)	499	(5.1)	**503**	(4.9)	**24.8**	(2.2)	6.2	(1.1)
	Liechtenstein	**515**	(10.6)	510	(12.8)	521	(11.7)	**571**	(10.8)	**24.1**	(4.9)	6.4	(2.6)
	Macao-China	522	(5.6)	522	(6.7)	540	(6.2)	546	(5.0)	**12.7**	(3.2)	1.7	(0.9)
	Russian Federation	**442**	(4.8)	489	(5.3)	491	(5.0)	**492**	(6.0)	**24.3**	(2.0)	4.3	(0.7)
	Serbia	**387**	(4.0)	414	(4.1)	437	(4.3)	**444**	(4.0)	**22.4**	(1.6)	7.4	(1.0)
	Thailand	**413**	(3.7)	418	(3.7)	428	(3.7)	**441**	(4.3)	**12.4**	(1.9)	2.0	(0.6)
	Tunisia	**334**	(3.7)	331	(3.2)	341	(3.8)	**374**	(4.1)	**22.5**	(2.4)	5.1	(1.0)
	Uruguay	**372**	(5.5)	399	(4.7)	423	(5.9)	**449**	(4.9)	**31.2**	(2.4)	6.9	(1.0)
	United Kingdom[1]	**482**	(3.4)	492	(3.7)	519	(4.2)	**545**	(4.3)	**24.7**	(1.7)	7.8	(1.0)

Note: Values that are statistically significant are indicated in bold (see Annex A4). "Min" is used for countries with more than 25 per cent of students at the lowest value on this index, which is –1.28. "Max" is used for countries with more than 25 per cent of students at the highest value of this index, which is 1.35.
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.5
Percentage of students and performance on the problem-solving scale, by type of family structure
Results based on students' self-reports

		Students from single-parent families				Students from other types of families				Difference in problem-solving performance (single-parent families – other types of families)	
		% of students	**S.E.**	**Mean score**	**S.E.**	**% of students**	**S.E.**	**Mean score**	**S.E.**	**Difference**	**S.E.**
OECD countries	Australia	20.0	(0.5)	511	(2.7)	80.0	(0.5)	535	(2.1)	**–25**	(2.5)
	Austria	15.9	(0.6)	509	(4.6)	84.1	(0.6)	508	(3.1)	1	(4.1)
	Belgium	17.0	(0.5)	496	(4.3)	83.0	(0.5)	535	(2.3)	**–39**	(3.9)
	Canada	18.6	(0.4)	514	(3.0)	81.4	(0.4)	538	(1.6)	**–24**	(3.0)
	Czech Republic	12.8	(0.5)	518	(3.8)	87.2	(0.5)	523	(3.2)	–5	(3.6)
	Denmark	24.3	(1.1)	499	(3.9)	75.7	(1.1)	523	(2.8)	**–25**	(3.8)
	Finland	20.0	(0.7)	543	(3.3)	80.1	(0.7)	549	(1.8)	–6	(3.0)
	France	20.3	(0.7)	510	(3.9)	79.8	(0.7)	523	(2.8)	**–14**	(3.9)
	Germany	16.7	(0.6)	514	(5.7)	83.3	(0.6)	521	(3.4)	–7	(5.1)
	Greece	23.4	(1.0)	430	(5.3)	76.6	(1.0)	455	(4.3)	**–26**	(5.0)
	Hungary	19.0	(0.7)	494	(4.3)	81.0	(0.7)	504	(2.9)	**–10**	(4.0)
	Iceland	13.3	(0.6)	496	(4.3)	86.7	(0.6)	507	(1.5)	**–11**	(4.6)
	Ireland	15.4	(0.7)	474	(4.1)	84.6	(0.7)	503	(2.4)	**–29**	(3.9)
	Italy	15.5	(0.6)	463	(4.6)	84.5	(0.6)	471	(3.1)	**–8**	(4.1)
	Japan	m	m	m	m	m	m	m	m	m	m
	Korea	20.3	(0.6)	548	(4.6)	79.7	(0.6)	551	(3.0)	–3	(3.4)
	Luxembourg	16.3	(0.5)	476	(3.9)	83.7	(0.5)	497	(1.5)	–21	(4.4)
	Mexico	33.1	(0.8)	378	(6.1)	66.9	(0.8)	389	(4.1)	**–10**	(4.3)
	Netherlands	13.7	(0.9)	500	(5.2)	86.3	(0.9)	527	(3.0)	**–28**	(5.1)
	New Zealand	18.9	(0.7)	515	(4.1)	81.1	(0.7)	538	(2.3)	**–22**	(4.4)
	Norway	27.1	(0.7)	473	(4.0)	72.9	(0.7)	498	(2.9)	**–25**	(4.2)
	Poland	11.4	(0.5)	479	(5.7)	88.7	(0.5)	488	(2.7)	–9	(5.2)
	Portugal	16.5	(0.6)	470	(5.9)	83.5	(0.6)	470	(3.9)	0	(4.7)
	Slovak Republic	11.5	(0.5)	487	(5.2)	88.5	(0.5)	494	(3.4)	–7	(4.6)
	Spain	14.0	(0.5)	475	(4.8)	86.0	(0.5)	484	(2.9)	–9	(4.8)
	Sweden	24.0	(0.7)	495	(3.6)	76.0	(0.7)	514	(2.5)	**–19**	(3.4)
	Switzerland	20.8	(0.7)	510	(3.9)	79.2	(0.7)	525	(3.3)	**–15**	(3.6)
	Turkey	32.7	(1.3)	402	(6.5)	67.3	(1.3)	412	(6.3)	**–10**	(4.4)
	United States	29.4	(0.9)	448	(4.2)	70.6	(0.9)	491	(3.0)	**–44**	(3.5)
	OECD total	23.4	(0.3)	458	(2.0)	76.6	(0.3)	492	(1.2)	–34	(1.6)
	OECD average	19.4	(0.1)	481	(1.1)	80.6	(0.1)	504	(0.6)	–23	(0.9)
Partner countries	Brazil	26.2	(0.9)	368	(6.8)	73.8	(0.9)	374	(4.7)	–6	(5.6)
	Hong Kong-China	19.7	(0.7)	534	(5.6)	80.3	(0.7)	552	(4.1)	**–18**	(4.0)
	Indonesia	9.9	(0.5)	347	(5.0)	90.2	(0.5)	363	(3.3)	**–16**	(4.5)
	Latvia	25.4	(0.9)	474	(5.1)	74.6	(0.9)	486	(4.0)	**–12**	(4.1)
	Liechtenstein	17.8	(2.1)	514	(12.1)	82.2	(2.1)	533	(4.8)	–18	(14.0)
	Macao-China	21.1	(1.3)	532	(6.7)	78.9	(1.3)	533	(2.9)	–1	(7.7)
	Russian Federation	20.7	(0.6)	479	(4.8)	79.3	(0.6)	481	(4.2)	–2	(3.4)
	Serbia	14.9	(0.7)	416	(4.6)	85.1	(0.7)	421	(3.4)	–6	(4.4)
	Thailand	21.7	(0.8)	420	(4.1)	78.3	(0.8)	427	(2.8)	**–7**	(3.6)
	Tunisia	7.3	(0.4)	332	(5.1)	92.7	(0.4)	348	(2.3)	**–16**	(5.2)
	Uruguay	23.1	(0.6)	409	(4.5)	76.9	(0.6)	412	(4.0)	–3	(4.3)
	United Kingdom[1]	22.2	(0.6)	494	(3.5)	77.8	(0.6)	515	(2.6)	**–21**	(3.3)

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.6
Percentage of students and performance on the problem-solving scale, by students' nationality and the nationality of their parents
Results based on students' self-reports

		Native students (born in the country of assessment with at least one of their parents born in the same country)				First-generation students (born in the country of assessment but whose parents were foreign-born)			
		% of students	S.E.	Mean score	S.E.	% of students	S.E.	Mean score	S.E.
OECD countries	Australia	77.3	(1.1)	535	(2.1)	11.7	(0.6)	521	(4.0)
	Austria	86.7	(1.0)	515	(3.2)	4.1	(0.5)	465	(9.9)
	Belgium	88.2	(0.9)	540	(2.4)	6.3	(0.6)	445	(7.7)
	Canada	79.9	(1.1)	535	(1.6)	9.2	(0.5)	532	(4.0)
	Czech Republic	98.7	(0.2)	523	(3.0)	0.5	(0.1)	c	c
	Denmark	93.5	(0.8)	522	(2.4)	3.5	(0.6)	443	(10.5)
	Finland	98.1	(0.2)	549	(1.8)	0.0	(0.0)	c	c
	France	85.7	(1.3)	529	(2.5)	10.8	(1.1)	482	(6.2)
	Germany	84.6	(1.1)	534	(3.4)	6.9	(0.8)	444	(9.2)
	Greece	92.6	(0.6)	452	(4.0)	0.5	(0.1)	c	c
	Hungary	97.7	(0.2)	502	(3.1)	0.1	(0.0)	c	c
	Iceland	99.0	(0.2)	507	(1.4)	0.2	(0.1)	c	c
	Ireland	96.5	(0.3)	499	(2.3)	1.0	(0.2)	c	c
	Italy	97.9	(0.3)	472	(3.0)	0.4	(0.1)	c	c
	Japan	99.9	(0.0)	548	(4.1)	0.0	(0.0)	c	c
	Korea	100.0	(0.0)	551	(3.1)	0.0	(0.0)	c	c
	Luxembourg	66.7	(0.6)	507	(1.8)	15.8	(0.6)	475	(3.7)
	Mexico	97.7	(0.3)	392	(4.3)	0.5	(0.1)	c	c
	Netherlands	89.0	(1.4)	532	(3.1)	7.1	(1.1)	463	(9.7)
	New Zealand	80.2	(1.1)	537	(2.5)	6.6	(0.7)	500	(7.5)
	Norway	94.4	(0.7)	495	(2.6)	2.3	(0.4)	452	(11.7)
	Poland	100.0	(0.0)	488	(2.7)	0.0	(0.0)	c	c
	Portugal	95.0	(1.4)	475	(3.3)	2.3	(0.4)	c	c
	Slovak Republic	99.1	(0.2)	493	(3.2)	0.6	(0.2)	c	c
	Spain	96.6	(0.4)	484	(2.7)	0.6	(0.1)	c	c
	Sweden	88.5	(0.9)	516	(2.2)	5.7	(0.5)	483	(8.9)
	Switzerland	80.0	(0.9)	538	(3.0)	8.9	(0.5)	480	(4.8)
	Turkey	99.0	(0.2)	409	(5.9)	0.5	(0.2)	c	c
	United States	85.6	(1.0)	483	(2.9)	8.3	(0.7)	465	(8.5)
	OECD total	*91.5*	*(0.3)*	*495*	*(1.1)*	*4.6*	*(0.2)*	*473*	*(4.5)*
	OECD average	*91.4*	*(0.2)*	*505*	*(0.6)*	*4.0*	*(0.1)*	*479*	*(2.0)*
Partner countries	Brazil	99.2	(0.2)	374	(4.7)	0.6	(0.2)	c	c
	Hong Kong-China	56.7	(1.4)	556	(4.1)	22.9	(0.9)	572	(4.0)
	Indonesia	99.7	(0.1)	364	(3.3)	0.2	(0.1)	c	c
	Latvia	90.6	(0.9)	483	(4.0)	8.3	(0.8)	487	(7.9)
	Liechtenstein	82.9	(2.0)	537	(4.5)	7.6	(1.3)	512	(17.9)
	Macao-China	23.9	(1.4)	536	(5.1)	57.9	(1.5)	533	(3.3)
	Russian Federation	86.5	(0.7)	482	(4.7)	6.4	(0.5)	473	(6.7)
	Serbia	91.1	(0.6)	423	(3.4)	3.2	(0.3)	417	(8.3)
	Thailand	99.9	(0.1)	426	(2.7)	0.1	(0.1)	c	c
	Tunisia	99.7	(0.1)	346	(2.1)	0.2	(0.1)	c	c
	Uruguay	99.2	(0.2)	411	(3.6)	0.4	(0.1)	c	c
	United Kingdom[1]	92.0	(0.8)	511	(2.4)	5.3	(0.6)	512	(7.3)

		Non-native students (foreign-born and whose parents were also foreign-born)				Difference in problem-solving performance between native and first-generation students		Difference in problem-solving performance between native and non-native students	
		% of students	S.E.	Mean score	S.E.	Difference	S.E.	Difference	S.E.
OECD countries	Australia	11.0	(0.7)	523	(4.8)	**14**	(4.3)	**12**	(4.7)
	Austria	9.2	(0.7)	453	(5.9)	**50**	(10.2)	**62**	(5.8)
	Belgium	5.5	(0.6)	446	(8.6)	**95**	(7.5)	**93**	(8.8)
	Canada	10.9	(0.8)	533	(4.7)	3	(4.2)	2	(4.7)
	Czech Republic	0.8	(0.1)	c	c	c	c	c	c
	Denmark	3.0	(0.4)	464	(8.8)	**79**	(10.5)	**58**	(8.7)
	Finland	1.8	(0.2)	c	c	c	c	c	c
	France	3.5	(0.5)	445	(14.8)	**47**	(6.5)	**84**	(14.9)
	Germany	8.5	(0.7)	461	(7.4)	**90**	(9.6)	**73**	(7.8)
	Greece	6.9	(0.7)	412	(7.0)	–13	(24.2)	**40**	(7.4)
	Hungary	2.2	(0.2)	c	c	c	c	c	c
	Iceland	0.8	(0.2)	c	c	c	c	c	c
	Ireland	2.5	(0.3)	c	c	c	c	c	c
	Italy	1.7	(0.2)	c	c	c	c	c	c
	Japan	0.1	(0.0)	c	c	c	c	c	c
	Korea	a	a	a	a	c	c	a	a
	Luxembourg	17.4	(0.5)	463	(3.9)	**33**	(4.2)	**44**	(4.4)
	Mexico	1.8	(0.2)	c	c	c	c	c	c
	Netherlands	3.9	(0.4)	462	(8.8)	**69**	(10.4)	**70**	(9.5)
	New Zealand	13.3	(0.7)	534	(4.6)	**38**	(8.1)	3	(5.3)
	Norway	3.4	(0.4)	417	(10.3)	**43**	(11.5)	**78**	(10.7)
	Poland	0.0	(0.0)	c	c	c	c	c	c
	Portugal	2.7	(1.1)	c	c	**c**	c	**c**	c
	Slovak Republic	0.3	(0.1)	c	c	c	c	c	c
	Spain	2.8	(0.4)	c	c	c	c	c	c
	Sweden	5.9	(0.7)	434	(10.1)	**33**	(8.3)	**82**	(10.4)
	Switzerland	11.1	(0.6)	447	(5.8)	**58**	(4.7)	**91**	(5.9)
	Turkey	0.5	(0.1)	c	c	c	c	c	c
	United States	6.1	(0.4)	446	(8.3)	**19**	(8.1)	37	(8.1)
	OECD total	*3.9*	*(0.1)*	*454*	*(3.9)*	*22*	*(4.4)*	*40*	*(4.0)*
	OECD average	*4.6*	*(0.1)*	*468*	*(1.9)*	*26*	*(2.0)*	*36*	*(1.9)*
Partner countries	Brazil	0.2	(0.1)	c	c	c	c	c	c
	Hong Kong-China	20.4	(1.3)	505	(5.0)	**–17**	(3.8)	**51**	(4.4)
	Indonesia	0.1	(0.0)	c	c	c	c	c	c
	Latvia	1.1	(0.2)	c	c	–5	(7.7)	c	c
	Liechtenstein	9.4	(1.6)	480	(19.6)	26	(18.6)	**58**	(20.7)
	Macao-China	18.2	(1.4)	531	(8.8)	4	(6.5)	6	(10.0)
	Russian Federation	7.0	(0.5)	451	(7.4)	9	(6.9)	**31**	(6.2)
	Serbia	5.6	(0.5)	423	(5.8)	6	(8.0)	–1	(5.5)
	Thailand	0.0	(0.0)	c	c	c	c	c	c
	Tunisia	0.1	(0.0)	c	c	c	c	c	c
	Uruguay	0.4	(0.1)	c	c	c	c	c	c
	United Kingdom[1]	2.7	(0.4)	c	c	0	(6.8)	c	c

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Table 5.7
Percentage of students and performance on the problem-solving scale, by language spoken at home
Results based on students' self-reports

		Language spoken at home most of the time IS DIFFERENT from the language of assessment, from other official languages or from other national dialects				Language spoken at home most of the time IS THE SAME as the language of assessment, other official languages or other national dialects				Difference in problem-solving performance (students speaking the same language – students speaking a different language)	
		% of students	S.E.	Mean score	S.E.	% of students	S.E.	Mean score	S.E.	Difference	S.E.
OECD countries	Australia	8.9	(0.7)	515	(5.3)	91.1	(0.7)	533	(1.9)	**18**	(5.2)
	Austria	9.0	(0.7)	458	(7.8)	91.0	(0.7)	514	(3.2)	**56**	(7.7)
	Belgium	4.8	(0.4)	450	(8.2)	95.2	(0.4)	539	(2.3)	**89**	(8.4)
	Canada	11.2	(0.7)	524	(4.4)	88.8	(0.7)	536	(1.6)	**11**	(4.3)
	Czech Republic	0.9	(0.2)	c	c	99.1	(0.2)	523	(3.1)	c	c
	Denmark	3.9	(0.5)	475	(10.2)	96.1	(0.5)	520	(2.5)	**44**	(10.4)
	Finland	1.8	(0.2)	c	c	98.2	(0.2)	549	(1.9)	c	c
	France	6.1	(0.7)	458	(9.1)	93.9	(0.7)	526	(2.5)	**69**	(9.3)
	Germany	7.7	(0.6)	430	(6.5)	92.3	(0.6)	531	(3.3)	**101**	(6.3)
	Greece	3.2	(0.4)	401	(11.0)	96.8	(0.4)	451	(3.9)	**49**	(11.0)
	Hungary	0.6	(0.1)	c	c	99.4	(0.1)	502	(3.0)	c	c
	Iceland	1.6	(0.2)	c	c	98.4	(0.2)	506	(1.4)	c	c
	Ireland	0.8	(0.2)	c	c	99.2	(0.2)	498	(2.3)	c	c
	Italy	1.6	(0.2)	c	c	98.4	(0.2)	474	(3.0)	c	c
	Japan	0.2	(0.1)	c	c	99.8	(0.1)	551	(4.1)	c	c
	Korea	0.1	(0.0)	c	c	99.9	(0.0)	551	(3.1)	c	c
	Luxembourg	25.0	(0.6)	464	(2.8)	75.0	(0.6)	507	(1.6)	**43**	(3.3)
	Mexico	1.1	(0.3)	c	c	98.9	(0.3)	386	(4.2)	c	c
	Netherlands	4.6	(0.6)	450	(10.0)	95.4	(0.6)	530	(3.1)	**79**	(10.3)
	New Zealand	9.0	(0.7)	516	(6.3)	91.0	(0.7)	535	(2.4)	**20**	(7.0)
	Norway	4.5	(0.5)	439	(9.7)	95.5	(0.5)	495	(2.5)	**56**	(9.8)
	Poland	0.2	(0.1)	c	c	99.8	(0.1)	487	(2.7)	c	c
	Portugal	1.4	(0.2)	c	c	98.6	(0.2)	472	(3.9)	c	c
	Slovak Republic	1.4	(0.3)	c	c	98.6	(0.3)	494	(3.2)	c	c
	Spain	1.7	(0.3)	c	c	98.3	(0.3)	482	(2.7)	c	c
	Sweden	6.9	(0.7)	456	(9.3)	93.1	(0.7)	516	(2.2)	**61**	(9.3)
	Switzerland	9.5	(0.7)	453	(6.7)	90.5	(0.7)	534	(3.4)	**81**	(6.3)
	Turkey	1.2	(0.6)	c	c	98.8	(0.6)	408	(6.0)	c	c
	United States	9.0	(0.7)	440	(7.2)	91.0	(0.7)	484	(3.0)	**44**	(7.1)
	OECD total	***4.5***	***(0.2)***	***449***	***(4.1)***	***90.7***	***(0.3)***	***494***	***(1.1)***	***46***	***(4.2)***
	OECD average	***4.5***	***(0.1)***	***465***	***(1.9)***	***91.2***	***(0.1)***	***504***	***(0.7)***	***39***	***(2.0)***
Partner countries	Brazil	0.5	(0.1)	c	c	99.5	(0.1)	372	(4.8)	c	c
	Hong Kong-China	4.5	(0.4)	473	(9.8)	95.6	(0.4)	553	(4.0)	**80**	(9.1)
	Indonesia	2.1	(0.3)	c	c	97.9	(0.3)	362	(3.3)	c	c
	Latvia	8.3	(1.1)	456	(8.8)	91.7	(1.1)	486	(3.9)	**30**	(8.1)
	Liechtenstein	18.4	(2.2)	505	(11.6)	81.6	(2.2)	543	(4.9)	**38**	(12.6)
	Macao-China	4.6	(0.7)	493	(13.6)	95.4	(0.7)	535	(2.8)	**42**	(14.6)
	Russian Federation	5.4	(1.3)	421	(13.3)	94.6	(1.3)	482	(4.4)	**61**	(12.4)
	Serbia	1.5	(0.2)	c	c	98.5	(0.2)	421	(3.4)	c	c
	Thailand	a	a	a	a	100.0	(0.0)	426	(2.7)	a	a
	Tunisia	0.4	(0.1)	c	c	99.6	(0.1)	344	(2.2)	c	c
	Uruguay	1.9	(0.4)	c	c	98.1	(0.4)	414	(3.7)	c	c
	United Kingdom[1]	3.8	(0.6)	489	(11.7)	96.2	(0.6)	512	(2.5)	**23**	(11.7)

Note: Values that are statistically significant are indicated in bold (see Annex A4).
1. Response rate too low to ensure comparability (see Annex A3).

Annex C

THE DEVELOPMENT AND IMPLEMENTATION OF PISA – A COLLABORATIVE EFFORT

Annex C: The development and implementation of PISA – a collaborative effort

Introduction

PISA is a collaborative effort, bringing together scientific expertise from the participating countries, steered jointly by their governments on the basis of shared, policy-driven interests.

A PISA Governing Board on which each country is represented determines, in the context of OECD objectives, the policy priorities for PISA and oversees adherence to these priorities during the implementation of the programme. This includes the setting of 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 that are charged with linking policy objectives with the best internationally available technical expertise. 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, the assessment materials have strong measurement properties, and the instruments place an emphasis on authenticity and educational validity.

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

The design and implementation of the surveys, within the framework established by the PISA Governing Board, is the responsibility of an international consortium, referred to as the PISA Consortium, led by the Australian Council for Educational Research (ACER). Other partners in this consortium include the Netherlands National Institute for Educational Measurement (Citogroep), The National Institute for Educational Research in Japan (NIER), the Educational Testing Service in the United States (ETS), and WESTAT in the United States.

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 PISA Governing Board, builds consensus among countries and serves as the interlocutor between the PISA Governing Board and the international consortium charged with the implementation of the activities. The OECD Secretariat also produces the indicators and analyses and prepares the international reports and publications in co-operation with the PISA consortium and in close consultation with Member countries both at the policy level (PISA Governing Board) and at the level of implementation (National Project Managers).

The following lists the members of the various PISA bodies and the individual experts and consultants who have contributed to PISA.

Members of the PISA Governing Board

Chair: Ryo Watanabe

Australia: Wendy Whitham

Austria: Helmut Bachmann and Jürgen Horschinegg

Belgium: Dominique Barthélémy, Christiane Blondin and Liselotte van de Perre

Brazil: Eliezer Pacheco

Canada: Satya Brink and Dianne Pennock

Czech Republic: Jan Koucky

Denmark: Jørgen Balling Rasmussen

Finland: Jari Rajanen

France: Gérard Bonnet

Germany: Hans Konrad Koch, Elfriede Ohrnberger and Botho Priebe

Greece: Vassilis Koulaidis

Hong Kong-China: Esther Ho Sui Chu

Hungary: Péter Vári

Iceland: Júlíus K. Björnsson

Indonesia: Bahrul Hayat

Ireland: Gerry Shiel

Italy: Giacomo Elias and Angela Vegliante

Japan: Ryo Watanabe

Korea: Kye Young Lee

Latvia: Andris Kangro

Luxembourg: Michel Lanners

Macao-China: Lam Fat Lo

Mexico: Felipe Martínez Rizo

Netherlands: Jules L. Peschar

New Zealand: Lynne Whitney

Norway: Alette Schreiner

Poland: Stanislaw Drzazdzewski

Portugal: Glória Ramalho

Russian Federation: Galina Kovalyova

Serbia: Dragica Pavlovic Babic

Slovak Republic: Vladimir Repas

Spain: Carme Amorós Basté, Guillermo Gil and Josu Sierra Orrantia

Sweden: Anita Wester

Switzerland: Katrin Holenstein and Heinz Rhyn

Thailand: Sunee Klainin

Tunisia: Néjib Ayed

Turkey: Sevki Karaca and Ruhi Kilç

United Kingdom: Lorna Bertrand and Liz Levy

United States: Mariann Lemke and Elois Scott

Uruguay: Pedro Ravela

Special Advisor: Eugene Owen

PISA 2003 National Project Managers

Australia: John Cresswell and Sue Thomson

Austria: Günter Haider and Claudia Reiter

Belgium: Luc van de Poele

Brazil: Mariana Migliari

Canada: Tamara Knighton and Dianne Pennock

Czech Republic: Jana Paleckova

Denmark: Jan Mejding

Finland: Jouni Välijärvi

France: Anne-Laure Monnier

Germany: Manfred Prenzel

Greece: Vassilia Hatzinikita

Hong Kong-China: Esther Ho Sui Chu

Hungary: Péter Vári

Iceland: Almar Midvik Halldorsson

Indonesia: Bahrul Hayat

Ireland: Judith Cosgrove

Italy: Maria Teresa Siniscalco

Japan: Ryo Watanabe

Korea: Mee-Kyeong Lee

Latvia: Andris Kangro

Luxembourg: Iris Blanke

Macao-China: Lam Fat Lo

Mexico: Rafael Vidal

Netherlands: Erna Gille

New Zealand: Fiona Sturrock

Norway: Marit Kjaernsli

Poland: Michal Federowicz

Portugal: Lídia Padinha

Russian Federation: Galina Kovalyova

Serbia: Dragica Pavlovic Babic

Slovak Republic: Paulina Korsnakova

Spain: Guillermo Gil

Sweden: Karin Taube

Switzerland: Huguette McCluskey

Thailand: Sunee Klainin

Tunisia: Néjib Ayed

Turkey: Sevki Karaca

United Kingdom: Rachael Harker, Graham Thorpe

United States: Mariann Lemke

Uruguay: Pedro Ravela

OECD Secretariat

Andreas Schleicher (overall co-ordination of PISA and Member country relations)

Miyako Ikeda (project management)

Claire Shewbridge (project management)

Claudia Tamassia (project management)

Sophie Vayssettes (statistical support)

Juliet Evans (administrative support)

Kate Lancaster (editorial support)

PISA Expert Groups

Mathematics Expert Group

Jan de Lange (Chair) (Utrecht University, The Netherlands)

Werner Blum (Chair) (University of Kassel, Germany)

Vladimir Burjan (National Institute for Education, Slovak Republic)

Sean Close (St Patrick's College, Ireland)

John Dossey (Consultant, United States of America)

Mary Lindquist (Columbus State University, United States of America)

Zbigniew Marciniak (Warsaw University, Poland)

Mogens Niss (Roskilde University, Denmark)

Kyung-Mee Park (Hongik University, Korea)

Luis Rico (University of Granada, Spain)

Yoshinori Shimizu (Tokyo Gakugei University, Japan)

Reading Expert Group

Irwin Kirsch (Chair) (Educational Testing Service, United States)

Marilyn Binkley (National Center for Educational Statistics, United States)

Alan Davies (University of Edinburgh, United Kingdom)

Stan Jones (Statistics Canada, Canada)

John de Jong (Language Testing Services, The Netherlands)

Dominique Lafontaine (Université de Liège Sart Tilman, Belgium)

Pirjo Linnakylä (University of Jyväskylä, Finland)

Martine Rémond (Institut National de Recherche Pédagogique, France)

Science Expert Group

Wynne Harlen (Chair) (University of Bristol, United Kingdom)

Peter Fensham (Monash University, Australia)

Raul Gagliardi (University of Geneva, Switzerland)

Svein Lie (University of Oslo, Norway)

Manfred Prenzel (Universität Kiel, Germany)

Senta A. Raizen (National Center for Improving Science Education (NCISE), United States)

Donghee Shin (KICE, Korea)

Elizabeth Stage (University of California, United States)

Problem Solving Expert Group

John Dossey (Chair) (Consultant, United States of America)

Beno Csapo (University of Szeged, Hungary)

Jan De Lange (Utrecht University, The Netherlands)

Eckhard Klieme (German Institute for International Educational Research, Germany)

Wynne Harlen (University of Bristol, United Kingdom)

Ton de Jong (University of Twente, The Netherlands)

Irwin Kirsch (Educational Training Service, United States)

Stella Vosniadou (University of Athens, Greece)

PISA Technical Advisory Group

Keith Rust (Chair) (Westat)

Ray Adams (ACER, Australia)

Pierre Foy (Statistics Canada, Canada)

Aletta Grisay (Belgium)

Larry Hedges (The University of Chicago, United States)

Eugene Johnson (American Institutes for Research, United States)

John de Jong (Language Testing Services, The Netherlands)

Irwin Kirsch (Educational Testing Service, United States)

Steve May (Ministry of Education, New Zealand)

Christian Monseur (HallStat SPRL, Belgium)

Norman Verhelst (Citogroep, The Netherlands)

J. Douglas Willms (University of New Brunswick, Canada)

PISA Consortium

Australian Council for Educational Research

Ray Adams (Project Director of the PISA Consortium)

Alla Berezner (data management, data analysis)

Eveline Gerbhardt (data processing, data analysis)

Marten Koomen (management)

Dulce Lay (data processing)

Le Tu Luc (data processing)

Greg Macaskill (data processing)

Barry McCrae (science instruments, test development mathematics and problem solving)

Martin Murphy (field operations and sampling)

Van Nguyen (data processing)

Alla Routitsky (data processing)

Wolfram Schulz (Coordinator questionnaire development. data processing, data analysis)

Ross Turner (Coordinator test development)

Maurice Walker (sampling, data processing, questionnaire development)

Margaret Wu (test development mathematics and problem solving, data analysis)

John Cresswell (test development science)

Juliette Mendelovits (test development reading)

Joy McQueen (test development reading)

Beatrice Halleux (translation quality control)

Westat

Nancy Caldwell (Director of the PISA Consortium for field operations and quality monitoring)

Ming Chen (weighting)

Fran Cohen (weighting)

Susan Fuss (weighting)

Brice Hart (weighting)

Sharon Hirabayashi (weighting)

Sheila Krawchuk (sampling and weighting)

Christian Monseur (consultant) (weighting)

Phu Nguyen (weighting)

Mats Nyfjall (weighting)

Merl Robinson (field operations and quality monitoring)

Keith Rust (Director of the PISA Consortium for sampling and weighting)

Leslie Wallace (weighting)

Erin Wilson (weighting)

Citogroep

Steven Bakker (science test development)

Bart Bossers (reading test development)

Truus Decker (mathematics test development)

Erna van Hest (reading test development and quality monitoring)

Kees Lagerwaard (mathematics test development)

Gerben van Lent (mathematics test development)

Ico de Roo (science test development)

Maria van Toor (office support and quality monitoring)

Norman Verhelst (technical advice, data analysis)

Educational Testing Service

Irwin Kirsch (reading test development)

Other experts

Cordula Artelt (questionnaire development)

Aletta Grisay (technical advice, data analysis, translation, questionnaire development)

Donald Hirsch (editorial review)

OECD PUBLICATIONS, 2 rue André-Pascal, PARIS CEDEX 16
PRINTED IN FRANCE
(962004131P1) ISBN 92-64-00642-7 – No. 53833 2004