3

A Question of Difficulty: Questions from PISA 2003

This chapter illustrates the PISA 2003 assessment with released assessment items and links to different levels of mathematical literacy proficiency. Actual assessment items can be found in this chapter, along with a discussion of students' performance on each of them.

INTRODUCTION

This chapter presents a number of characteristics of the PISA questions in relation to different levels of proficiency in mathematics. The characteristics discussed include the proficiency descriptions used to report the different levels of performance of students in the PISA mathematics assessment and the related issue of how difficult the question is, the type of response required by students (*e.g.* select a given response or write a short answer), and the role of the context that the question is set in. The complexity of the language used and other aspects of the presentation of questions will also be discussed.

DESCRIBING GROWTH IN MATHEMATICAL LITERACY: HOW DIFFICULT IS THE QUESTION AND WHERE DOES IT FIT ON THE PISA MATHEMATICS SCALE?

The questions used in the PISA mathematics assessment cover a wide range of difficulties. This is necessary in order to obtain valid and reliable ability estimates for the range of students sampled in different countries. The difficulty of the questions used can be illustrated by reference to the PISA mathematics scale that was developed to quantify performance in different countries (OECD, 2004a). This chapter discusses factors that contribute to the difficulty of questions in PISA mathematics.

PISA mathematics questions cover a wide range of difficulties in a wide range of formats.

The PISA mathematics questions take a variety of formats, and while Chapter 5 analyses more extensively the relationship between the type of question and how difficult the question is, the basic types of PISA mathematics questions are briefly introduced here. In the 2003 assessment, all mathematics questions broadly either required students to construct a response or to select a response. In the case of the latter, these could be either simple multiple-choice questions, requiring the students to select one answer from a number of optional responses, or complex multiple-choice questions, presenting students with a small number of statements and requiring students to select from given optional responses for each statement (such as “true” or “false”). In the case of questions where students need to construct a response, this could be either an extended response (*e.g.* extensive writing, showing a calculation, or providing an explanation or justification of the solution given) or a short answer (*e.g.* a single numeric answer, or a single word or short phrase; and sometimes a slightly more extended short response). Much of the discussion around reform in mathematics education involves questions presented in context and requiring communication as part of the response (de Lange, 2007). The analyses of item difficulty in this chapter and, later, in Chapters 4 and 5, focuses on how questions were presented to students and the degree to which students were able to meet the challenges posed by the items.

The methodology of the PISA assessment, including the sampling design, the design of the assessment instruments including the various types of questions, and the methods used to analyse the resulting data, leads to efficient estimates of the proportion of students in each country lying at various parts of the *mathematical literacy* scale. *Mathematical literacy* is conceived as a continuous variable, and the scale has been developed to quantify and describe this. The PISA *mathematical literacy*

scale is constructed to have a mean of 500 score points and a standard deviation of 100 score points; that is, about two-thirds of the 15-year-olds across OECD countries score between 400 and 600 score points. Six proficiency levels are defined for the *mathematical literacy* scale, and the kinds of student behaviours typical in each of those proficiency levels are described. This "described proficiency scale" is central to the way in which PISA reports comparative performance in mathematics.

The difficulty of PISA mathematics questions is determined using three different approaches…

… with simple percentages …

This report uses three different but related methods to quantify and refer to the difficulty of the mathematics items. First, the simplest approach involves using "percent correct" data (that is, the percentage of students in each country or internationally correctly answering a question). This form of comparison is useful when the focus is on an individual question (for example, comparing the success rate of male and female students on a particular item, or of students in different countries on a particular item), or on comparing the performance on two questions by a particular group of students.

… logistic models …

Second, the formal statistical analysis of PISA data is carried out using units called a "logit". A logit represents the logarithm of the ratio of the probability of a correct answer to an item to the probability of an incorrect answer [often called the log-odds ratio]. For example an item with a probability correct of 0.50 would have a logit value equal to 0 [log (0.5/0.5) = log (1) = 0]. The use of log-odds ratio transforms the infinite scale associated with the probability ratio through logarithms to a 0-1 scale estimation of the location of the difficulty of all items and the ability of all students on a single dimension. Item performance can then be placed on a single scale by their log-odds ratio. This approach is basic to item response theory and its depiction of item difficulty and other item parameters through varied parameter models using the logistic function. In particular, the use of logit scores for items places them on a linear scale allowing for arithmetic computations with the logit unit (Thissen & Wainer, 2001). This is useful in comparing the relative strengths and weaknesses of items, and students on the PISA mathematics framework within each country and is discussed extensively in Chapter 4. Third, the "logits" units for each question are transformed as described to form the PISA mathematics scale giving an associated score in PISA score points. This approach allows for each PISA mathematics question to be located along the same scale and thus shows the relative difficulty of each question.

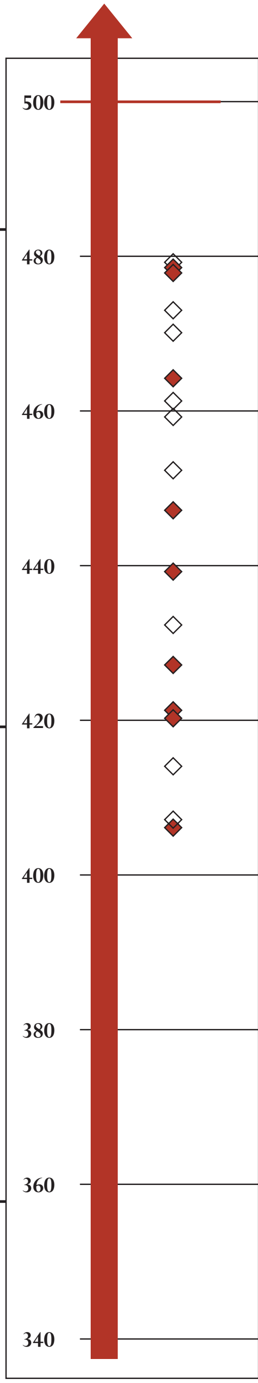
… and the statistically calculated PISA mathematics scale.

THE PISA SCALE AND DIFFICULTY

Figures 3.1a, 3.1b and 3.1c illustrate the placement of items on the PISA mathematics examination in terms of their relation to the PISA scale's scores. Released items and student performance on them are illustrated in the following sections in an explanation of student performance associated with various intervals on the PISA mathematics scale.

PISA releases some questions after the assessment to help illustrate the kind of mathematics problems that students have to solve. Thirty-one of the mathematics questions used in the PISA 2003 assessment were publicly released and

Figure 3.1a ■ **PISA mathematics proficiency Levels 1 and 2: Competencies students typically show and publicly released questions**



OECD average performance in mathematics

Proficiency Level 2
(420 to 482 score points)

At Level 2 students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results.

Sixteen questions are at this proficiency level of which seven are released.

CUBES – Question 1
GROWING UP – Question 1
SKATEBOARD – Question 1 (*Partial credit*)
THE BEST CAR – Question 1
EXCHANGE RATE – Question 2
EXPORTS – Question 1
STAIRCASE – Question 1

Proficiency Level 1
(358 to 420 score points)

At Level 1 students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli.

Four questions are at this proficiency level of which two are released.

GROWING UP – Question 2 (*Partial credit*)
EXCHANGE RATE – Question 1

Below Level 1

Two questions are below proficiency Level 1.

Twenty questions with scores at these levels.

Figure 3.1b ■ **PISA mathematics proficiency Levels 3 and 4: Competencies students typically show and publicly released questions**

Proficiency Level 4 (544 to 607 score points)

At Level 4 students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic ones, linking them directly to aspects of real-world situations. Students at this level can utilise well-developed skills and reason flexibly, with some insight, in these contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments, and actions.

Twenty-six questions are at this level of which twelve are released.

Proficiency Level 3 (482 to 544 score points)

At Level 3 students can execute clearly described procedures, including those that require sequential decisions. They can select and apply simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They can develop short communications reporting their interpretations, results and reasoning

Seventeen questions are at this level of which six are released.

Forty-three questions with scores at these levels, of which eighteen are released.

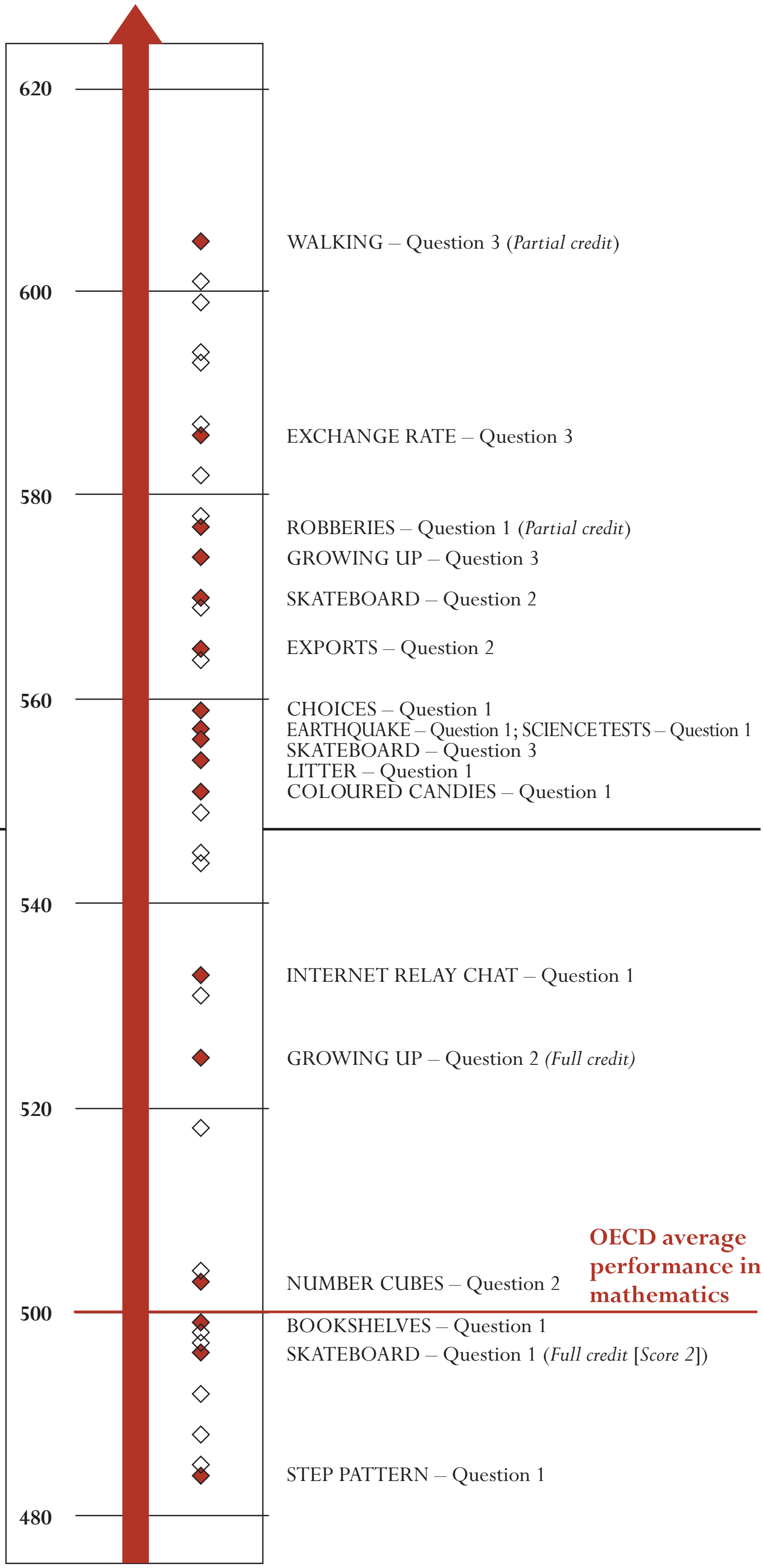


Figure 3.1c ■ **PISA mathematics proficiency Levels 5 and 6: Competencies students typically show and publicly released questions**

Proficiency Level 6
(669 score points and above)

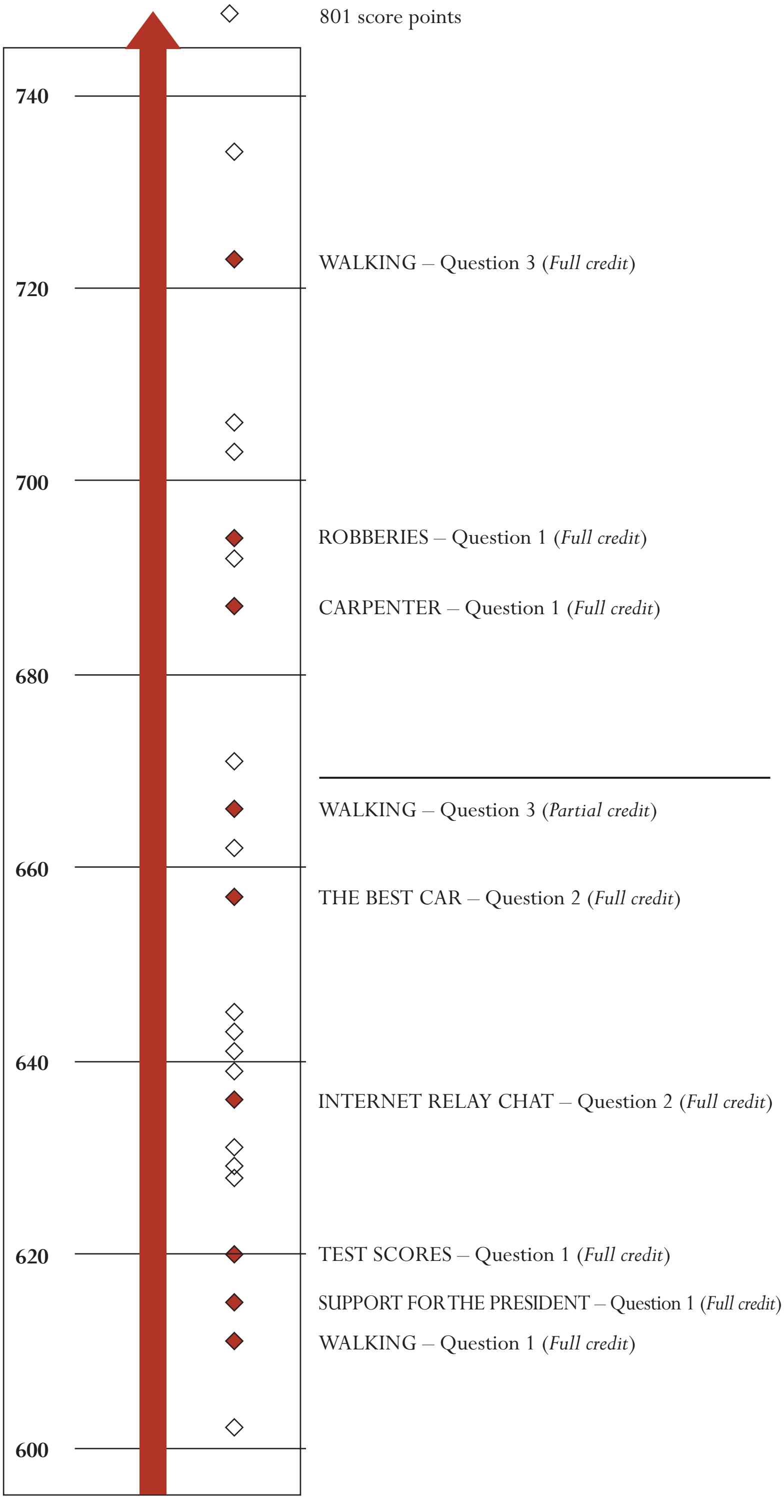
Students can conceptualise, generalise, and utilise information based on their investigations and modelling of complex problem situations. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understandings along with a mastery of symbolic and formal mathematical operations and relationships to develop new approaches and strategies for attacking novel situations. Students at this level can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situations.

Ten questions are at this level of which three are released.

Proficiency Level 5
(607 to 669 score points)

At Level 5 students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They can reflect on their actions and formulate and communicate their interpretations and reasoning.

Seventeen questions are at this level of which six are released.



Twenty-seven questions with scores at these levels.

Figures 3.1a, 3.1b and 3.1c show where each of these questions is located on the PISA *mathematical literacy* scale. It is useful to remember that the OECD average performance in PISA 2003 mathematics is 500 score points. Most of the questions in Figures 3.1a, 3.1b and 3.1c involve simple scoring, where credit is awarded only if the answer is correct and a 0 is awarded otherwise. However, five of these questions involve the use of up to three different scoring categories. For these questions, the term "full credit" is used to describe a fully correct answer, and one or more "partial credit" categories exist for answers that are only partially correct, for example the student may have only solved the first step of the problem at hand or have shown all necessary working, but made a minor calculation error. As a result, for the 31 questions in Figures 3.1a, 3.1b and 3.1c result in a total of 36 different scores as shown in Figures 3.1a, 3.1b, and 3.1c. Student performance through these score levels helps illustrate the full range of PISA mathematics proficiency (Levels 1 to 6, where Level 1 is the simplest and Level 6 the hardest). Figure 3.1a shows the summary descriptions of what students can typically do at PISA mathematics proficiency Levels 1 and 2 where the easiest questions in the mathematics assessment are located. The PISA score points for all questions included in Levels 1 and 2 are below the OECD average performance of 500 score points. They range from 358 to 482 score points.

Student performance is measured on a scale with an average score of 500. Students are grouped in six levels of proficiency, plus a group below Level 1.

The remainder of the chapter presents the 31 released PISA 2003 mathematics questions to illustrate more fully the different levels of proficiency in mathematics and to analyse the characteristics related to the difficulty of the question. Questions are presented in three distinct sections: the easiest questions in PISA 2003 mathematics illustrating PISA proficiency at Levels 1 and 2 (in fact the two easiest questions in the test lie below Level 1) which are found on the PISA scale from 358 to 482 points; questions of moderate difficulty in PISA 2003 mathematics illustrating proficiency at Levels 3 and 4, which are found on the PISA scale from 482 to 607 points; and the most difficult questions in PISA 2003 mathematics illustrating proficiency at Levels 5 and 6 which are found on the PISA scale from 607 and above. In each section an introductory summary table presents the following key characteristics for all questions: the associated PISA score points on the *mathematical literacy* scale (including, where appropriate, scores for both full and partial credit); where the question fits into the three main components of the PISA mathematics framework – content area or "overarching idea", competency cluster and context; the format used for the question; the traditional mathematics topic tested most prominently in the question; and the length of question (as measured by a simple word count) to indicate the reading demand. Additional information and data on the test items and related student performance can be found at *www.pisa2003.acer.edu.au/downloads.php* at the Australian Council for Educational Research's PISA website for PISA 2003.

EXAMPLES OF THE EASIEST MATHEMATICS QUESTIONS FROM PISA 2003

In the PISA 2003 mathematics assessment, the two easiest questions lie below Level 1, and 20 questions are included in proficiency Levels 1 and 2. Nine of these 20 questions are released (coming from seven units) and these are listed in Table 3.1 along with the difficulty of each question on the PISA scale and other characteristics. Several of these questions are included in units that contain more than one question. In the case that these questions are located at different proficiency Levels (*e.g.* the unit EXPORTS contains Question 1 at Level 2 and Question 2 at Level 4) each question is presented at the section of the chapter related to where its score points appear on the PISA scale. Figures showing country level performances on many of these items are found in Annex A1, Figures A1.1 through A1.8.

Recall from Figure 3.1a that Level 1 proficiency indicates that students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli. Level 2 students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results.

Table 3.1 shows that there is a prevalence of questions in the *reproduction* competency cluster among the easiest questions in PISA 2003 mathematics. Fourteen of the 20 questions at Levels 1 and 2 are in the *reproduction* competency cluster and this is also true of the two easiest questions lying below Level 1. In general, questions in the *reproduction* competency cluster place lower-level cognitive demands on students, and are therefore easier. Nevertheless there are relatively easy questions also from the *connections* competency cluster (six of the 20 questions in Levels 1 and 2 are in this category). All four content areas are represented among the easier questions in the PISA 2003 mathematics assessment: seven questions belong to *quantity*, six to *change and relationships*, five to *space and shape* and two to *uncertainty*. However, all nine released items for Levels 1 and 2 employed the short response item format.

Each of the items listed in Figure 3.1a as appearing in Levels 1 and 2 is now examined in detail, with performance information for students from the participating countries used as a lens to understand both student work and differences between the countries. In addition to the presentation of each unit in the right-hand side of the following displays, the scoring guide with sample responses for each level is presented beneath the questions contained within each question within the unit. Additional information about the items and about technical aspects associated with the scaling of the scores can be found in the international report (OECD, 2004a) and the technical report detailing the operational aspects of the PISA 2003 study (OECD, 2005).

Table 3.1

Characteristics of the easiest released PISA 2003 mathematics questions

Item code	Question	OECD average percent correct	Location on PISA scale (PISA score points): Question	Full/Partial credit points: 1	2	3	Traditional topic	Content area ("Overarching Idea")	Competency cluster	Context ('Situation')	Length of question[1]	Response format
M413Q01	EXCHANGE RATE – *Question 1*	80	406	**406**			Number	Quantity	Reproduction	Public	Medium	Short Answer
M150Q02[2]	GROWING UP – *Question 2* Partial 1 Point	69	472	**420**	525		Data	Change and relationships	Reproduction	Scientific	Medium	Short Answer
M547Q01	STAIRCASE – *Question 1*	78	421	**421**			Number	Space and shape	Reproduction	Educational and occupational	Short	Short Answer
M438Q01	EXPORTS – *Question 1*	79	427	**427**			Data	Uncertainty	Reproduction	Public	Medium	Short Answer
M413Q02	EXCHANGE RATE – *Question 2*	74	439	**439**			Number	Quantity	Reproduction	Public	Short	Short Answer
M704Q01	THE BEST CAR – *Question 1*	73	447	**447**			Algebra	Change and relationships	Reproduction	Public	Long	Short Answer
M520Q01	SKATEBOARD – *Question 1* Partial 1 Point	72	480	**464**	496		Number	Quantity	Reproduction	Personal	Long	Short Answer
M150Q01	GROWING UP – *Question 1*	67	477	**477**			Number	Change and relationships	Reproduction	Scientific	Short	Short Answer
M145Q01	CUBES – *Question 1*	68	478	**478**			Data	Space and shape	Reproduction	Educational and occupational	Medium	Short Answer

1. Short questions contain fewer than 50 words. Medium-length questions contain 51 to 100 words. Long questions contain more than 100 words; word count in relation to question difficulty is discussed in detail in Chapter 5.
2. Note items with partial credit are presented where that partial credit fits and the full discussion is contained later in the chapter at the level where full credit for the item is obtained. The bold type in the score column indicates the score being represented in a given problem line.

EXCHANGE RATE

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

Question 1: EXCHANGE RATE

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

1 SGD = 4.2 ZAR

Mei-Ling changed 3000 Singapore dollars into South African rand at this exchange rate.

How much money in South African rand did Mei-Ling get?

Answer: ....

EXCHANGE RATE – Question 1 was the third easiest question of all the PISA 2003 mathematics questions. On average across OECD countries, about 80% of students solved this problem correctly.

Context: *Public* – currency exchange associated with international travel

Content area: *Quantity* – quantitative relationships with money

Competency cluster: *Reproduction*

The question requires students to:

- Interpret a simple and explicit mathematical relationship.
- Identify and carry out the appropriate multiplication.
- Reproduce a well-practised routine procedure.

Students were most successful on this question in the partner country Liechtenstein (95%), the partner economy Macao-China (93%), Finland (90%), and France and the partner economy Hong Kong-China (89%). Most students attempted to answer this question, with only 7% failing to respond, on average, across OECD countries.

Question 2: EXCHANGE RATE

On returning to Singapore after 3 months, Mei-Ling had 3 900 ZAR left. She changed this back to Singapore dollars, noting that the exchange rate had changed to:

1 SGD = 4.0 ZAR

How much money in Singapore dollars did Mei-Ling get?

Answer: ....

EXCHANGE RATE – Question 2 is slightly more difficult, but still among the easiest of the PISA 2003 mathematics questions. On average across OECD countries, about 74% of students were able to do this successfully. This Level 2 item had a PISA difficulty level of 439.

Context: *Public* – currency exchange associated with international travel

Content area: *Quantity* – quantitative relationships with money

Competency cluster: *Reproduction*

The question requires students to:

- Recognise the change in the context from Question 1 that results from the need to convert money in the "opposite direction".
- Carry out a division to find the required answer.

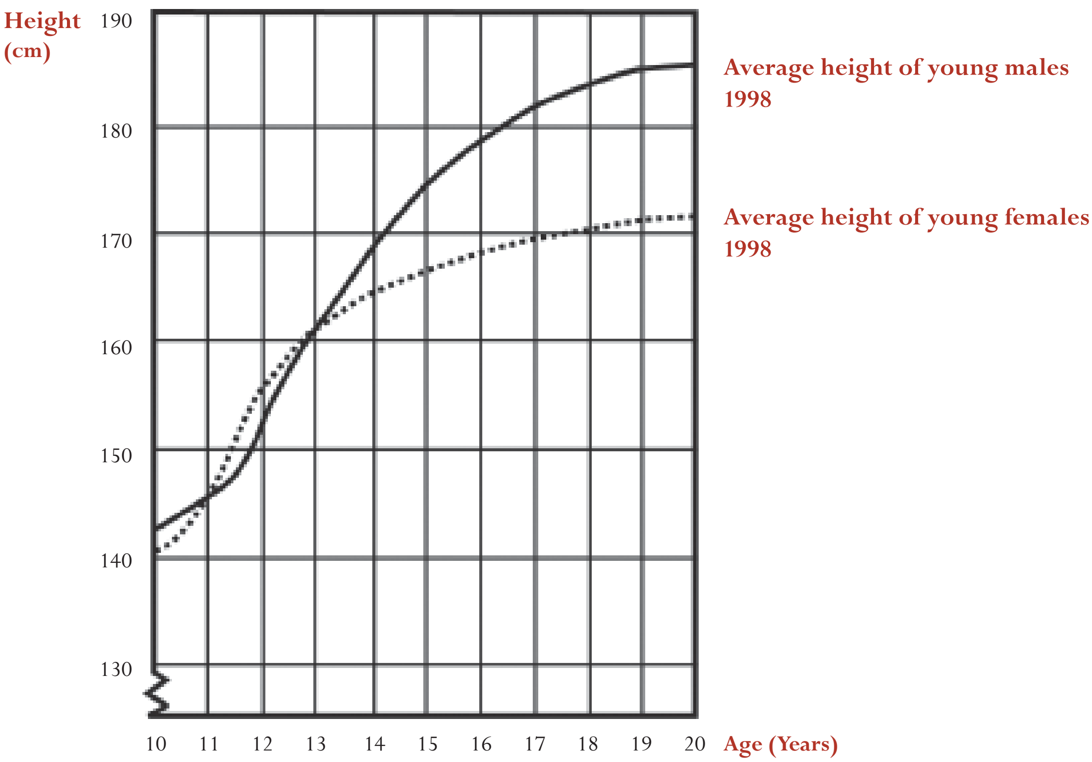
Students were most successful on this question in the partner country Liechtenstein (93%), the partner economy Macao-China (89%), Finland and the partner economy Hong Kong-China (88%), Austria (87%), and France, Switzerland and the Slovak Republic (85%). There was a 9% non-response rate across the OECD countries as a whole, while 14-17% of students in Turkey, Italy, Norway, Portugal, Greece and the partner countries Uruguay and Tunisia, failed to respond. In the partner country Brazil, 27% of students failed to respond. This compares to less than 5% of students in Finland, Canada, the Netherlands and the partner countries/economies Macao-China, Liechtenstein and Hong Kong-China.

Note that Question 3 of this unit is presented in the section *Examples of moderate to difficult questions in the PISA 2003 mathematics assessment*.

GROWING UP

Youth grows taller

In 1998 the average height of both young males and young females in the Netherlands is represented in this graph.



Question 1: GROWING UP

Since 1980 the average height of 20-year-old females has increased by 2.3 cm, to 170.6 cm.

What was the average height of a 20-year-old female in 1980?

Answer: .... cm

GROWING UP – Question 1 illustrates Level 2 in PISA 2003 mathematics and has a difficulty of 477 PISA score points. On average across OECD countries, 67% of students were able to do this successfully.

Context: *Scientific* – the growth curves of young males and females over a period of ten years. Science is no different from the real world in the sense that it uses graphical representation frequently, for example the graph in this question representing changes in height in relation to age.

Content area: *Change and relationships* – focus on change in height in relation to age. Basic mathematical operation of subtraction.

Competency cluster: *Reproduction* – basic thinking and reasoning involving the most basic questions (How much is the difference?); basic argumentation where the student just needs to follow a standard quantitative process. There is some added complexity in the fact that the answer can be found by ignoring the graph altogether – an example of redundant information.

The question requires students to:

- Extract the relevant information from a single source (and ignore the graph which is a redundant source).
- Make use of a single representational mode.
- Employ a basic subtraction algorithm (170.6 – 2.3).

Students were most successful on this question in Korea (82%), France (80%), Japan and the partner country the Russian Federation (78%), Sweden and Iceland (76%), the Czech Republic (75%) and the Slovak Republic (74%).

Most students attempted to answer this question – only 8% failed to do so across OECD countries and this concerned less than 1% of students in the Netherlands. However, 23% of students in Greece and 21% of students in the partner country Serbia did not respond to this question.

Question 2: GROWING UP

According to this graph, on average, during which period in their life are females taller than males of the same age?

....

GROWING UP – Question 2 illustrates two different levels of proficiency depending on whether students gave a fully or partially correct answer. Here, a partially correct answer scored at 1 point illustrates exactly the boundary between Level 1 and Level 2 with a difficulty of 420 PISA score points. A fully correct answer illustrates Level 3 with a difficulty of 525 score points. On average across OECD countries, 28% of students were only capable of achieving the partial 1 point level.

Context: *Scientific*

Content area: *Change and relationships* – focus on the relationship between age and height. The mathematical content can be described as belonging to the "data" domain: the students are asked to compare characteristics of two data sets, interpret these data sets and draw conclusions.

Competency cluster: *Reproduction* – interpret and decode reasonably familiar and standard representations of well known mathematical objects. Students need to think and reason (where do the graphs have common points?), use argumentation to explain which role these points play in finding the desired answer and communicate and explain the argumentation. However, all these competencies essentially involve reproduction of practised knowledge.

The question requires students to:

- Interpret and use a graph. Make conclusions directly from a graph. Report the results of their reasoning in a precise manner.

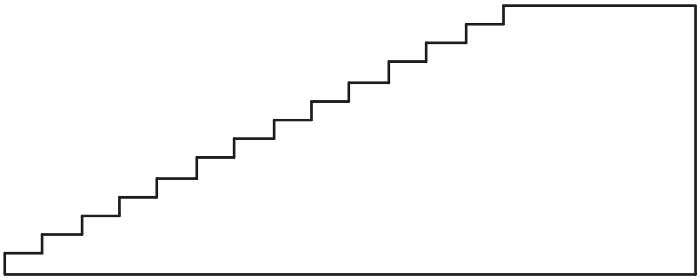
Students were considered to give a partially correct answer if they properly identified ages like 11 and/or 12 and/or 13 as being part of the answer, but failed to identify the continuum from 11 to 13 years. These students were able to compare the two graphs properly, but did not communicate their answer adequately or failed to show sufficient insight into the fact that the answer would be an interval. This is probably in part due to the fact that the proper procedure may not have been routine. On average across the OECD countries, 28% of students gave a partially correct answer showing that their reasoning and/or insight was well directed, but failed to come up with a full, comprehensive answer. This was the case for 43% of students in the United States, 42% of students in the Slovak Republic and the partner country Thailand, 40% of students in Poland, and between 39 and 37% of students in Italy, the Czech Republic, Sweden and the partner countries/economies the Russian Federation and Macao-China.

Seven percent of students on average across the OECD countries did not attempt to answer this question. This concerned less than 3% of students in the Netherlands, Finland, Canada and the partner economy Macao-China.

Note that the discussion of full credit for GROWING UP Q1 will be presented in the discussion of Level 3 questions in the section *Examples of moderate to difficult questions in the PISA 2003 mathematics assessment*.

STAIRCASE

Question 1: STAIRCASE



The diagram above illustrates a staircase with 14 steps and a total height of 252 cm:

What is the height of each of the 14 steps?

Height: ……………… cm.

STAIRCASE – Question 1 illustrates Level 2 in PISA 2003 mathematics, with a difficulty of 421 PISA score points (just one point over the boundary of Level 1 and 2). On average across OECD countries, 78% of students were able to do this successfully.

Context: *Educational and occupational* – situated in a daily life context for carpenters (for example). One does not need to be a carpenter to understand the relevant information; it is clear that an informed citizen should be able to interpret and solve a problem like this that uses two different representation modes: language, including numbers, and a graphical representation. But the illustration serves a simple and non-essential function: students know what stairs look like. This item is noteworthy because it has redundant information (the depth is 400 cm) which is sometimes considered by students as confusing but a common feature in real-world problem solving.

Content area: *Space and shape* – graphical representation of a staircase, but the actual procedure to carry out is a simple division.

Competency cluster: *Reproduction* – carry out a basic operation. Students solve the problem by invoking and using standard approaches and procedures in one way only. All the required information, and even more than required, is presented in a recognisable situation.

The question requires students to:

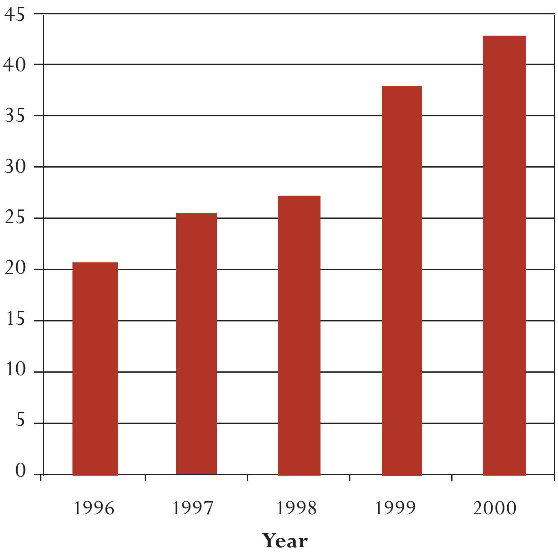
- Extract the relevant information from a single source.
- Apply of a basic algorithm (divide 252 by 14).

In each OECD country the majority of students gave the correct answer "18", but this was especially true in the partner economy Macao-China (89%), the partner economy Hong Kong-China (87%), Switzerland (86%), Finland, the Netherlands and the partner country Liechtenstein (85%). On average across OECD countries 10% of students did not respond to this question. However in Hungary 29% of students did not respond the question, as did 25-26% of students in the partner countries Indonesia, Brazil and Thailand.

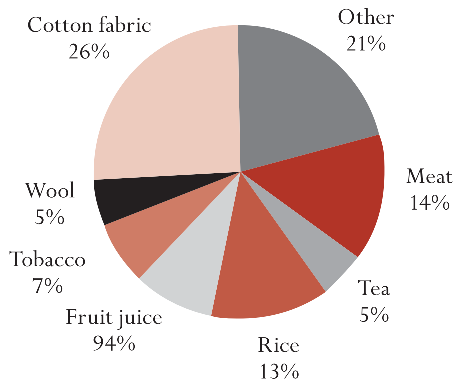
EXPORTS

The graphics below show information about exports from Zedland, a country that uses zeds as its currency.

Total annual exports from Zedland in millions of zeds, 1996-2000



Distribution of exports from Zedland in 2000



Question 1: EXPORTS

What was the total value (in millions of zeds) of exports from Zedland in 1998?

Answer: ....

EXPORTS – Question 1 illustrates Level 2 in PISA 2003 mathematics, with a difficulty of 427 PISA score points. On average across OECD countries, 79% of students were able to do this successfully.

Context: *Public* – The information society in which we live relies heavily on data, and data are often represented in graphics. The media use graphics often to illustrate articles and make points more convincingly. Reading and understanding this kind of information therefore is an essential component of *mathematical literacy*.

Content area: *Uncertainty* – the focus is on exploratory data analysis. The mathematical content is restricted to reading data from a bar graph or pie chart.

Competency cluster: *Reproduction* – interpret and recognise situations in contexts that require no more than direct inference. Students need to solve the problem by decoding and interpreting a familiar, practised standard representation of a well known mathematical object.

The question requires students to:

- Follow the written instructions.
- Decide which of the two graphs is relevant.
- Locate the correct information in that graph.

Successful students answered either “27.1 million zeds” or “27 100 000 zeds” or even just “27.1” without the unit “zeds”. Students were most successful on this question in France (92%), the Netherlands (91%), Canada (90%), the partner country Liechtenstein (89%), Belgium, Portugal and Finland (88%). On average across OECD countries only 7% of students failed to respond to this question.

Note that there is another question in this unit (EXPORTS – Question 2) and this is presented in the section *Examples of moderate to difficult questions in PISA 2003 mathematics*.

THE BEST CAR

A car magazine uses a rating system to evaluate new cars, and gives the award of "The Car of the Year" to the car with the highest total score. Five new cars are being evaluated, and their ratings are shown in the table.

Car	Safety Features (S)	Fuel Efficiency (F)	External Appearance (E)	Internal Fittings (T)
Ca	3	1	2	3
M2	2	2	2	2
Sp	3	1	3	2
N1	1	3	3	3
KK	3	2	3	2

The ratings are interpreted as follows:

3 points = Excellent

2 points = Good

1 point = Fair

Question 1: THE BEST CAR

To calculate the total score for a car, the car magazine uses the following rule, which is a weighted sum of the individual score points:

$$\text{Total Score} = (3 \times S) + F + E + T$$

Calculate the total score for Car "Ca". Write your answer in the space below.

Total score for "Ca": ....

THE BEST CAR – Question 1 illustrates Level 2 in PISA 2003 mathematics, with a difficulty of 447 PISA score points. On average across OECD countries, 73% of students were able to do this successfully.

Context: *Public* – an article in a car magazine is a very familiar context, especially for males. The underlying mathematics is relevant for males and females as everyone is presented with this kind of problem, that is, the evaluation of a consumer good using a rating system, whether it be cars, washing machines, coffee makers, etc. This is therefore an important part of *mathematical literacy*.

Content area: *Change and relationships* – the focus is on the relationship of numbers in a formula

Competency cluster: *Reproduction* – students need to reproduce a practised procedure. However, this is not trivial as it involves an equation and students typically find it difficult to work with equations presented in such a real-world context.

The question requires students to:

- Read and understand a relatively straightforward question.
- Multiply a number by 3.
- Add four simple numbers.

Successful students answered "15 points". Students were most successful on this question in the partner economy Macao-China (90%), the partner countries/economies Liechtenstein and Hong Kong-China (87%), Korea (84%), Canada (82%) and Denmark, Austria and Japan (80%). On average across OECD countries, 10% of students did not respond to this question, but this included 24% of students in Mexico, 23% in the partner country Brazil and 21% of students in Greece. It is interested to note that the OECD average for female students was 74.5%, while that for male students was 71.33% – a significant difference favouring female students!

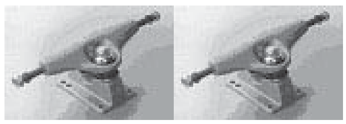
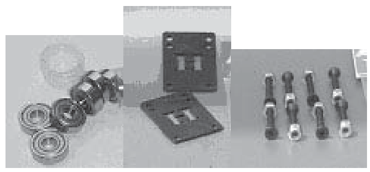
Note that there is another question in this unit (THE BEST CAR – Question 2) and this is presented in the section *Examples of difficult questions in PISA 2003 mathematics*.

SKATEBOARD

Eric is a great skateboard fan. He visits a shop named SKATERS to check some prices.

At this shop you can buy a complete board. Or you can buy a deck, a set of 4 wheels, a set of 2 trucks and a set of hardware, and assemble your own board.

The prices for the shop's products are:

Product	**Price in zeds**	
Complete skateboard	82 or 84	
Deck	40, 60 or 65	
One set of 4 Wheels	14 or 36	
One set of 2 Trucks	16	
One set of hardware (bearings, rubber pads, bolts and nuts)	10 or 20	

Question 1: SKATEBOARD

Eric wants to assemble his own skateboard. What is the minimum price and the maximum price in this shop for self-assembled skateboards?

(a) Minimum price: .... zeds.

(b) Maximum price: .... zeds.

SKATEBOARD – Question 1 illustrates two different levels of proficiency depending on whether students gave a fully or partially correct answer. Here, a partially correct answer scored at 1 point illustrates a Level 2 performance with a difficulty of 464 PISA score points. A fully correct answer illustrates Level 3 with a difficulty of 496 PISA score points. On average across OECD countries, 11% of students were only capable of achieving the partial 1 point level.

Context: *Personal* – skateboards are part of the youth culture; either students skateboard themselves or watch others do it – especially on television.

Content area: *Quantity* – the students are asked to find a minimum and maximum price for the construction of a skateboard, under given numerical conditions. The skills needed to solve this problem are certainly an important part of *mathematical literacy* as they make it possible to make more informed decisions in daily life.

Competency cluster: *Reproduction* – students need to solve the problem by finding a simple strategy to produce the minimum and maximum and reproduce practised knowledge in combination with the performance of a routine addition.

The question requires students to:

- Interpret the question correctly and so understand that they need to provide two answers.
- Extract the relevant information from a simple table.
- Find a simple strategy to come up with the minimum and maximum (this is simple because the strategy that seems trivial actually works: for the minimum take the lower numbers, for the maximum the larger ones).
- Perform a basic addition. (The whole number addition: 40 + 14 + 16 + 10 equals 80, gives the minimum, and the maximum is found by adding the larger numbers: 65 + 36 + 16 + 20 = 137).

This question illustrates Level 2 when the students give a partially correct answer: by giving either the minimum or the maximum, but not both. On average across the OECD countries 11% of students gave a partially correct answer. This was the case for 28% of students in France and 13% of students in Mexico, Luxembourg and the partner country Serbia.

On average across the OECD countries, the majority of students responded answer this question – with only 5% failing to do so. (Although this was 12% of students in Turkey and 11% of students in Greece and Japan).

Note that full credit for SKATEBOARD Q1 displays mathematical proficiency at Level 3 is discussed for this performance in the section *Examples of moderate to difficult questions in PISA 2003 mathematics*.

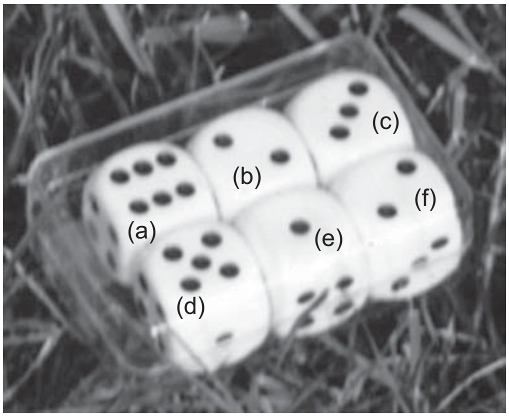
CUBES

Question 1: CUBES

In this photograph you see six dice, labelled (a) to (f). For all dice there is a rule:

The total number of dots on two opposite faces of each die is always seven.

Write in each box the number of dots on the ***bottom*** *face of the dice corresponding to the photograph.*



(a)	(b)	(c)

(d) (e) (f)

CUBES – Question 1 illustrates Level 2 in PISA 2003 mathematics with a difficulty of 478 PISA score points. On average across OECD countries, 68% of students were able to do this successfully.

Context: *Educational and occupational* – the context in this case is somewhat difficult to classify: for many students the number cubes are very familiar recreational objects and therefore the context could be classified as "personal", but these objects are also frequently seen in school contexts. Further, the question calls for spatial representation skills that are required at a basic level in many occupations.

Content area: *Space and shape* – spatial representation

Competency cluster: *Reproduction* – apply a simple given rule and use basic spatial representation skills. Even if students are not familiar with number cubes (or dice) the essential rule is stated clearly in the introductory text. These competencies are essential to mathematical literacy, but are only required a very basic level here.

The question requires students to:

- Apply the rule, that the opposites sum up to 7, six times.
- Use spatial representation skills to 'transfer' the presented photo into the table.

Successful students answered "Top row (1 5 4) Bottom Row (2 6 5)" or drew a diagram showing the correct numbers on the faces of the cubes. Students were most successful on this question in Finland and Switzerland (80% correct), Japan (79%), Sweden (78%) and France, Canada and the partner country Liechtenstein (76%). Most students across the OECD countries responded to this question – only 6% failed to do so, although this was 12% of students in Hungary, Greece and in the partner countries Serbia, Tunisia and Brazil.

So, what characteristics do these examples of easier mathematics questions share beyond the predominance of questions in the *reproduction* competency cluster? First, the response formats used in this set of easier questions are similar. All eight questions require students to undertake rather convergent thinking and provide a simple, short and rather closed constructed response, usually a single numeric answer. None of these released questions requires students to write an explanation of their solution, or a justification of their result. All of the released easy questions involve rather directed instructions towards finding a single correct numeric answer, with little reasoning required. In general, it is observed that questions of these forms are the easiest to answer. They place no demands on students in relation to deciding what kind of response would constitute an answer to the question asked. Students can find or calculate the answer, or they cannot. In most cases, the question formats even indicated where or how the student should respond.

Second, there is a low level of complexity in the language used in questions and the unit or stimulus they are related to. Most of the easy questions have relatively low reading demands. The text in the stimulus of each unit generally consists of simple, direct statements. Similarly, the questions are relatively short and direct. The language demand of questions can be an important factor in determining the difficulty of questions and is discussed in more detail in Chapter 5. Finally, the graphical or pictorial displays in the setting of the units are also of familiar formats and ones that students would have had experience in creating or manipulating in school or life situations.

EXAMPLES OF MODERATE TO DIFFICULT MATHEMATICS QUESTIONS FROM PISA 2003

Forty-three of the mathematics questions used in the PISA 2003 assessment lie on that part of the *mathematical literacy* scale covered by Levels 3 and 4, and can therefore be regarded as representing intermediate levels of difficulty. Eighteen of these questions, coming from 15 different units, have been released. These are listed together with information about various characteristics of the questions in Table 3.2. Figures showing country level performances on many of these moderate to difficult items are found in Annex A1, Figures A.1.9 through A.1.23.

Around this middle part of the reported PISA *mathematical literacy* scale, the increased difficulty of questions relative to those in Levels 1 and 2 can be seen in the increased number of questions from the *connections* competency cluster, and the appearance of questions from the *reflection* competency cluster. Mathematics questions in these competency clusters typically have greater complexity and impose increased cognitive demands. All four of the mathematical content areas are represented among the released questions from these levels. Similarly, all contexts are represented. And unlike the easier questions discussed earlier, there is a mix of different response formats. Almost half of these released questions are of the short answer type, similar to the entire set of easier questions, but there are also multiple-choice questions and questions requiring an extended response. It is also evident that these questions impose a greater reading load than those in Levels 1 and 2.

Table 3.2
Characteristics of moderate to difficult released questions

Item Code	Question	OECD average percent correct	Location on PISA scale (PISA score points): Question	Full/partial credit points: 1	2	3	Traditional topic	Content area ("Overarching Idea")	Competency cluster	Context ("Situation")	Length of question[1]	Response format
M806Q01	STEP PATTERN – *Question 1*	66	484	**484**			Number	Quantity	Reproduction	Educational and occupational	Short	Short Answer
M520Q01	SKATEBOARD – *Question 1*	72	480	464	**496**		Number	Quantity	Reproduction	Personal	Long	Short Answer
M484Q01	BOOKSHELVES – *Question 1*	61	499	**499**			Number	Quantity	Connections	Educational and occupational	Medium	Short Answer
M555Q02	NUMBER CUBES – *Question 2*	63	503	**503**			Geometry	Space and shape	Connections	Personal	Long	Complex Multiple Choice
M150Q02[2]	GROWING UP – *Question 2*	69	472	420	**525**		Data	Change and relationships	Reproduction	Scientific	Medium	Short Answer
M402Q01	INTERNET RELAY CHAT – *Question 1*	54	533	**533**			Measurement	Change and relationships	Connections	Personal	Medium	Short Answer
M467Q01	COLOURED CANDIES – *Question 1*	50	549	**549**			Data	Uncertainty	Reproduction	Personal	Short	Multiple Choice
M505Q01	LITTER – *Question 1*	52	551	**551**			Data	Uncertainty	Reflection	Scientific	Medium	Extended Response
M520Q03	SKATEBOARD – *Question 3*	50	554	**554**			Number	Quantity	Connections	Personal	Long	Short Answer
M468Q01	SCIENCE TESTS – *Question 1*	47	556	**556**			Data	Uncertainty	Reproduction	Educational and occupational	Medium	Short Answer
M509Q01	EARTHQUAKE – *Question 1*	46	557	**557**			Data	Uncertainty	Reflection	Scientific	Long	Multiple Choice
M510Q01	CHOICES – *Question 1*	49	559	**559**			Number	Quantity	Connections	Educational and occupational	Medium	Short Answer

Table 3.2
Characteristics of moderate to difficult released questions *(continued)*

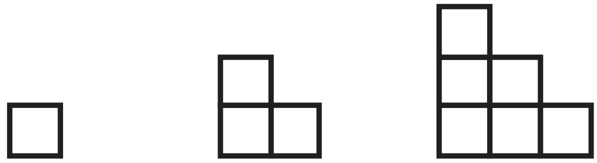
Item Code	Question	OECD average percent correct	Location on PISA scale (PISA score points): Question	Full/partial credit points: 1	2	3	Traditional topic	Content area ("Overarching Idea")	Competency cluster	Context ("Situation")	Length of question[1]	Response format
M438Q02	EXPORTS – *Question 2*	48	565	**565**			Number	*Uncertainty*	*Connections*	*Public*	Medium	Multiple Choice
M520Q02	SKATEBOARD – *Question 2*	46	570	**570**			Number	*Quantity*	*Reproduction*	*Personal*	Short	Multiple Choice
M150Q03	GROWING UP – *Question 3*	45	574	**574**			Data	*Change and relationships*	*Connections*	*Scientific*	Medium	Extended Response
M179Q01	ROBBERIES – *Question 1* Partial 1 Point	30	635	**577**	694		Data	*Uncertainty*	*Connections*	*Public*	Short	Extended Response
M413Q03	EXCHANGE RATE – *Question 3*	40	586	**586**			Number	*Quantity*	*Reflection*	*Public*	Medium	Extended Response
M124Q03	WALKING – *Question 3* Partial 1 point	21	665	**605**	666	723	Algebra	*Change and relationships*	*Connections*	*Personal*	Medium	Extended Response

1. Short questions contain fewer than 50 words. Medium-length questions contain 51 to 100 words. Long questions contain more than 100 words. Length of question in relation to question difficulty is discussed in detail in Chapter 5.

STEP PATTERN

Question 1: STEP PATTERN

Robert builds a step pattern using squares. Here are the stages he follows.



Stage 1 **Stage 2** **Stage 3**

As you can see, he uses one square for Stage 1, three squares for Stage 2 and six for Stage 3.

How many squares should he use for the fourth stage?

Answer: .... squares.

STEP PATTERN – Question 1 illustrates Level 3 in PISA 2003 mathematics, with a difficulty of 484 PISA score points. On average across OECD countries, 66% of students were able to do this successfully.

Context: *Educational and occupational* – This problem would be representative of similar tasks seen commonly in mathematics classes or textbooks; indeed it is almost a pure mathematical problem. The importance of such a question in a test focusing on *mathematical literacy* is not immediately clear. Such questions are not seen in newspapers, on television, or at work. But recognising regularities or patterns and being able to predict the next member of the sequence are helpful skills when structured processing is required. It is a well known fact that problems like these appear in many psychological tests. Some mathematicians do not approve of questions that ask for the next member of a given string of integers, as it can be argued mathematically that any answer is correct. For students of this age this turns out, in practice, not to be a problem and certainly is not for this particular question since it presents a pattern with both a numeric and geometric base.

Content area: *Quantity* – recognising a pattern from a numeric and geometric base.

Competency cluster: *Reproduction* – use of very basic strategies and no need for mathematisation. The question is simple and clearly stated and it is not strictly necessary to read the text. There are at least two simple possible strategies: Either count the numbers of each object (1, 3, 6 and the next number will be 10); or sketch the next object and then count the number of squares.

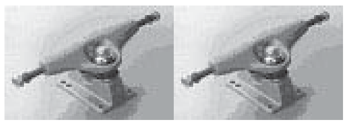
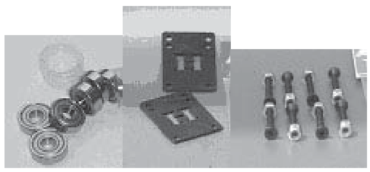
Successful students answered "10". Students were most successful on this question in Japan (88%), the partner economy Hong Kong-China (83%), Korea and the partner economy Macao-China (80%) and the Czech Republic, Denmark and Norway (78%). This was a question that the majority of students responded to – on average across OECD countries only 1% failed to do so and this did not surpass 5% in any of the OECD countries.

SKATEBOARD

Eric is a great skateboard fan. He visits a shop named SKATERS to check some prices.

At this shop you can buy a complete board. Or you can buy a deck, a set of 4 wheels, a set of 2 trucks and a set of hardware, and assemble your own board.

The prices for the shop's products are:

Product	Price in zeds	
Complete skateboard	82 or 84	
Deck	40, 60 or 65	
One set of 4 Wheels	14 or 36	
One set of 2 Trucks	16	
One set of hardware (bearings, rubber pads, bolts and nuts)	10 or 20	

Question 1: SKATEBOARD

Eric wants to assemble his own skateboard. What is the minimum price and the maximum price in this shop for self-assembled skateboards?

(a) Minimum price: .... zeds.

(b) Maximum price: .... zeds.

SKATEBOARD – Question 1 illustrates two different levels of proficiency depending on whether students gave a fully or partially correct answer. A partially correct answer illustrates Level 2 and has been discussed earlier. A fully correct answer illustrates Level 3 with a difficulty of 496 PISA score points. On average across OECD countries, 72% of students were able to do this successfully.

Context: *Personal* – skateboards are part of the youth culture; either students skateboard themselves or watch others do it – especially on television.

Content area: *Quantity* – the students are asked to find a minimum and maximum price for the construction of a skateboard, under given numerical conditions. The skills needed to solve this problem are certainly an important part of *mathematical literacy* as they make it possible to make more informed decisions in daily life.

Competency cluster: *Reproduction* – students need to solve the problem by finding a simple strategy to produce the minimum and maximum and reproduce practised knowledge in combination with the performance of a routine addition.

The question requires students to:

- Interpret the question correctly and so understand that they need to provide two answers.
- Extract the relevant information from a simple table.
- Find a simple strategy to come up with the minimum and maximum (this is simple because the strategy that seems trivial actually works: for the minimum take the lower numbers, for the maximum the larger ones).
- Perform a basic addition. (The whole number addition: 40 + 14 + 16 + 10 equals 80, gives the minimum, and the maximum is found by adding the larger numbers: 65 + 36 + 16 + 20 = 137).

Students were most successful on this question, providing both the minimum (80) and the maximum (137), in Finland (81%), the partner country Liechtenstein and Switzerland (76%), Canada (75%), Australia, New Zealand, Belgium and Austria (74%).

On average across the OECD countries, the majority of students responded to this question – with only 5% failing to do so (although this was 12% of students in Turkey and 11% of students in Greece and Japan).

Question 2: SKATEBOARD

The shop offers three different decks, two different sets of wheels and two different sets of hardware. There is only one choice for a set of trucks.

How many different skateboards can Eric construct?

A 6

B 8

C 10

D 12

SKATEBOARD – Question 2 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 570 score points. On average across OECD countries, 46% of students were able to do this successfully.

Context: *Personal*

Content area: *Quantity* – routine computation. The skills needed to solve this problem are certainly an important part of *mathematical literacy* as they make it possible to make more informed decisions in daily life.

Competency cluster: *Reproduction* – all the required information is explicitly presented. Students need to understand what the required "strategy" is and then carry out that strategy. For students having identified the required strategy, the mathematics involves the basic routine computation: 3 x 2 x 2 x 1. However, if students do not have experience with such combinatorial calculations, their strategy might involve a systematic listing of the possible combinatorial outcomes. There are well-known algorithms for this (such as a tree diagram). The strategy to find the number of combinations can be considered as common, and more or less routine. It involves following and justifying standard quantitative processes, including computational processes, statements and results.

The question requires students to:

- Interpret correctly a text in combination with a table.
- Apply accurately a simple enumeration algorithm.

Successful students answered "D" (12). Students were most successful on this question in Japan (67%), Korea (65%), Denmark and the partner economy Hong Kong-China (60%).

The incorrect answer most frequently given by students across the OECD countries was "A" (25%), followed by "B" (18%). Only in Korea and Hungary did more students chose answer "B" than "A", and in Japan and the Netherlands students were equally divided among these two incorrect categories. To get answer "B" students may have added the whole numbers in the question to get a total of 8. To get answer "A" it is most likely that students misread the question and missed one of the components with two different sets (either the wheels or the hardware).

Question 3: SKATEBOARD

Eric has 120 zeds to spend and wants to buy the most expensive skateboard he can afford.

How much money can Eric afford to spend on each of the 4 parts? Put your answer in the table below.

Part	Amount (zeds)
Deck	
Wheels	
Trucks	
Hardware	

SKATEBOARD – Question 3 illustrates the lower part of Level 4, with a difficulty of 554 PISA score points (ten points above the boundary with Level 3). On average across OECD countries, 50% of students were able to do this successfully.

Context: *Personal*

Content area: *Quantity* – students are asked to compute what is the most expensive Skateboard that can be bought for 120 zeds by using some kind of quantitative process that has not been described. But this task is certainly not straightforward: there is no standard procedure or routine algorithm available.

Competency cluster: *Connections* – students need to use an independent and not routine problem-solving approach. Students may use different strategies in order to find the solution, including trial and error. So the setting of this problem is familiar or quasi-familiar but the problem to be solved is not simply routine. Students have to pose a question (How do we find…?), look at the table with prices, make combinations and do some computation. A strategy that will work with this problem is to first use all the higher values, and then adjust the answer, reducing the price until the desired maximum of 120 zed is reached. So taking the deck at 65 zed, the wheels at 36 zed, the trucks at 16 zed (no choice here) and the hardware at 20 zed. This gives a total of 137 zed – the maximum found earlier in Question 1. The cost needs to be reduced by at least 17 zed. It is possible to reduce the cost by 5 zed or 25 zed on the deck, by 22 zed on the wheels or by 10 zed on the hardware. The best solution is clear: save 22 zed on the wheels.

The question requires students to:

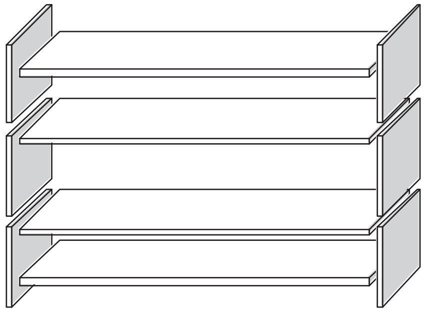
- Reason in a familiar context.
- Connect the question with the data given in the table, or in other words, relate text-based information to a table representation.
- Apply a non-standard strategy.
- Carry out routine calculations.

Successful students answered "65 zeds on a deck, 14 on wheels, 16 on trucks and 20 on hardware". Students were most successful on this question in the partner economy Macao-China (65%), the partner economy Hong Kong-China (62%), Finland, Canada and Sweden (59%), Belgium (58%) and Australia (57%). Seventeen percent of students across the OECD countries narrowly missed the correct answer and only gave correct prices for three of the four parts and this was as much as 27% of students in the partner country Thailand, 26% in Mexico, 21% in the partner country Serbia, and 20% in Luxembourg, the United States and Greece.

BOOKSHELVES

Question 1: BOOKSHELVES

To complete one set of bookshelves a carpenter needs the following components:



4 long wooden panels,

6 short wooden panels,

12 small clips,

2 large clips and

14 screws.

*The carpenter has in stock 26 long wooden panels,
33 short wooden panels, 200 small clips, 20 large clips and 510 screws.*

How many sets of bookshelves can the carpenter make?

Answer: ....

BOOKSHELVES – Question 1 illustrates Level 3 in PISA 2003 mathematics, with a difficulty of 499 PISA score points. On average across the OECD countries, 61% of students were able to do this successfully.

Context: *Educational and occupational* – The stem uses both a visualisation as well as text, with a lot of numbers, and a clear and short question. Almost by definition problems from an occupational context fit well with *mathematical literacy*. The problem in principle also has a certain level of authenticity: it stands for large collection of problems that have as core the attribution of parts to a production process in order to optimise the quantities of required components and to minimise waste.

Content area: *Quantity* – computation of ratios. Students compute the following ratio for each of the components: available components/required components per set of bookshelves. This gives: 26/4 (for long panels); 33/6 (for short panels); 200/12 (for small clips); 20/2 (for large clips); 510/14 (for screws).

Competency cluster: *Connections* – strategic thinking and some mathematisation. Students analyse the ratios they have computed to find that the smallest answer is 33/6, or 5.5. However, this is only an indication of the solution to the problem and 5.5 would not be a satisfactory response to the question asked. Students need to interpret this mathematical answer back into the bookshelves context to find the correct real-world solution: 5 sets of bookshelves.

The question requires students to:

- Develop a strategy to connect two bits of information for each component: the number available, and the number needed per set of bookshelves.
- Use logical reasoning to link that analysis across the components to produce the required solution.
- Communicate the mathematical answer as a real-world solution.

Successful students answered "5". Students were most successful on this question in Finland and the partner economy Hong Kong-China (74%), Korea, the Czech Republic, Belgium and Denmark (72%). On average across OECD countries 29% of students responded to this question but gave an incorrect answer and 10% did not respond at all.

NUMBER CUBES

Question 2: NUMBER CUBES

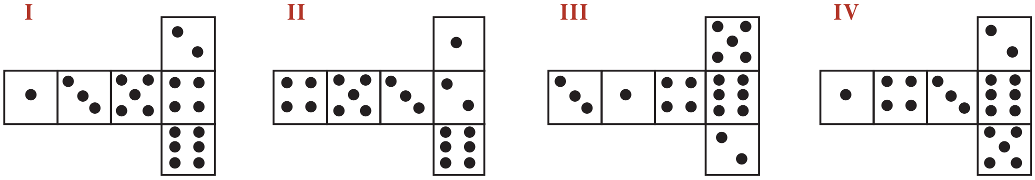
On the right, there is a picture of two dice.

Dice are special number cubes for which the following rule applies:

The total number of dots on two opposite faces is always seven.

You can make a simple number cube by cutting, folding and gluing cardboard. This can be done in many ways. In the figure below you can see four cuttings that can be used to make cubes, with dots on the sides.

Which of the following shapes can be folded together to form a cube that obeys the rule that the sum of opposite faces is 7? For each shape, circle either "Yes" or "No" in the table below.



Shape	Obeys the rule that the sum of opposite faces is 7?
I	Yes / No
II	Yes / No
III	Yes / No
IV	Yes / No

NUMBER CUBES – Question 2 illustrates Level 3 in PISA 2003 mathematics, with a difficulty of 503 PISA score points. On average across OECD countries, 63% of students were able to do this successfully.

Context: *Personal* – many games that children encounter during their education, whether formal or informal, use number cubes. The problem does not assume any previous knowledge about this cube, in particular the rule of construction: two opposite sides have a total of seven dots.

Content area: *Space and shape* – spatial reasoning skills. The given construction rule emphasises a numerical aspect, but the problem posed requires some kind of spatial insight or mental visualisation technique. These competencies are an essential part of *mathematical literacy* as students live in three-dimensional space, and often are confronted with two-dimensional representations. Students need to mentally imagine the four plans of number cubes reconstructed into a three-dimensional number cube and judge whether they really obey the numerical construction rule.

Competency cluster: *Connections* – The problem is certainly not routine: students need to connect written information, graphical representation and interpret back-and-forth. However, all the relevant information is clearly presented in writing and with graphics.

The question requires students to:

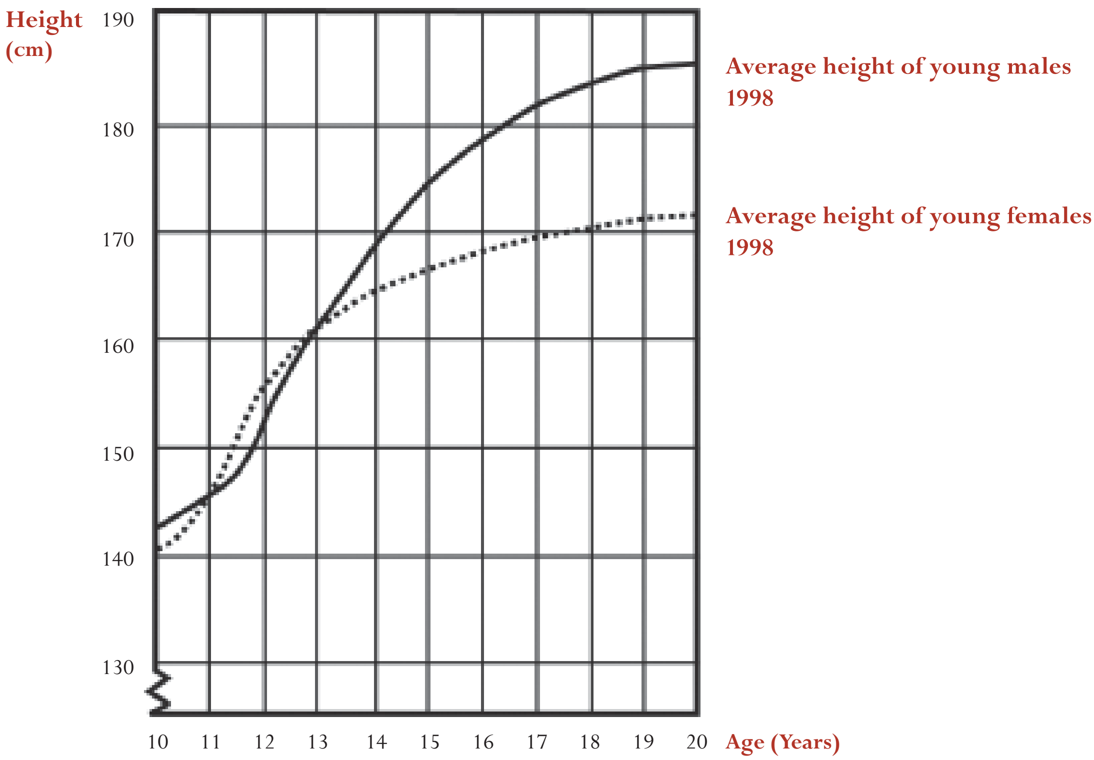
- Encode and interpret spatially two-dimensional objects.
- Interpret the connected three-dimensional objects.
- Interpret back-and-forth between model and reality.
- Check certain basic quantitative relations.

Successful students answered "No, Yes, Yes, No" in that order. Students were most successful on this question in Japan (83% correct), Korea (81%), Finland (76%), Belgium (74%), the Czech Republic and Switzerland (73%). A further 16% of students on average across OECD countries narrowly missed the fully correct answer and provided three out of four of the correct shapes and only 2% of students did not respond to the question.

GROWING UP

Youth grows taller

In 1998 the average height of both young males and young females in the Netherlands is represented in this graph:



Question 2: GROWING UP

According to this graph, on average, during which period in their life are females taller than males of the same age?

....

GROWING UP – Question 2 illustrates two different levels of proficiency depending on whether students gave a fully or partially correct answer. A partially correct answer illustrates exactly the boundary between Level 1 and Level 2 with a difficulty of 420 PISA score points. Here a fully correct answer illustrates Level 3 with a score of 2 points for PISA scale difficult of 525 score points. On average across OECD countries, 69% of students were able to do this successfully.

Context: *Scientific*

Content area: *Change and relationships* – focus on the relationship between age and height. The mathematical content can be described as belonging to the "data" domain: the students are asked to compare characteristics of two data sets, interpret these data sets and draw conclusions.

Competency cluster: *Reproduction* – interpret and decode reasonably familiar and standard representations of well known mathematical objects. Students need to think and reason (where do the graphs have common points?), use argumentation to explain which role these points play in finding the desired answer and communicate and explain the argumentation. However, all these competencies essentially involve reproduction of practised knowledge.

The question requires students to:

- Interpret and use a graph.
- Make conclusions directly from a graph.
- Report the results of their reasoning in a precise manner.

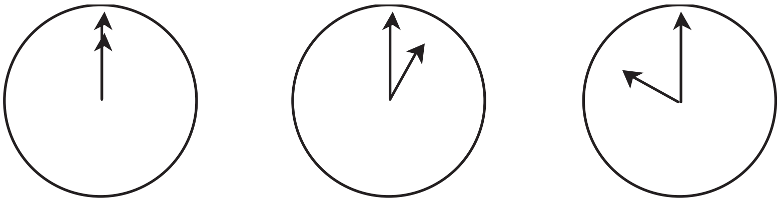
Students who were most successful on this question showed that their reasoning and/or insight was well directed and properly identified the continuum from 11 to 13 years. This was the case for 80% of students in Korea, 74% in the partner country Liechtenstein, 72% in France, 69% in Belgium and 67% in the Netherlands and Finland. Across countries, the majority of successful students communicated the correct interval as follows: "Between age 11 and 13"; "From 11 years old to 13 years old, girls are taller than boys on average"; or "11-13". However, a minority of successful students stated the actual years when girls are taller than boys, which is correct in daily-life language: "Girls are taller than boys when they are 11 and 12 years old"; or "11 and 12 years old." This concerned only 5% or less of the fully correct answers given in 16 of the OECD countries and this only surpassed 10% of fully correct answers in Turkey (21%), Mexico (19%) and Ireland (13%).

Seven percent of students on average across the OECD countries did not respond to this question. This concerned less than 3% of students in the Netherlands, Finland, Canada and the partner economy Macao-China.

INTERNET RELAY CHAT

Mark (from Sydney, Australia) and Hans (from Berlin, Germany) often communicate with each other using "chat" on the Internet. They have to log on to the Internet at the same time to be able to chat.

To find a suitable time to chat, Mark looked up a chart of world times and found the following:



Greenwich 12 Midnight   Berlin 1:00 AM   Sydney 10:00 AM

Question 1: INTERNET RELAY CHAT

At 7:00 PM in Sydney, what time is it in Berlin?

Answer: ....

INTERNET RELAY CHAT – Question 1 illustrates Level 4 with a difficulty of 533 PISA score points. On average across OECD countries 54% of students were able to do this successfully.

Context: *Personal* – this assumes either that students are familiar at some level with chatting over the internet, and/or they know about time differences in this or another context.

Content area: *Change and relationships*

Competency cluster: *Connections* – solving a non-routine problem, using simple mathematical tools, and making use of different representations. The problem does need some mathematisation, starting with identifying the relevant mathematics. The question is simple, and so are the numbers and the actual operations needed (adding and subtracting whole numbers). So the complexity lies really in the mathematisation: first the students have to identify the time difference between Berlin and Sydney (9 hours). Then they have to appreciate the fact that it is 9 hours later in Sydney. Then they have to apply this difference to the new situation.

This question requires students to:

- Identify the relevant mathematics.
- Solve a non-routine, but simple problem.
- Use different representations.

Successful students answered "10 AM or 10:00". At least 60% of students answered this question correctly in the Czech Republic, Denmark, France, Belgium, Germany, Switzerland, Luxembourg, Korea, Japan, the Slovak Republic and Austria, as well as in the partner country Liechtenstein.

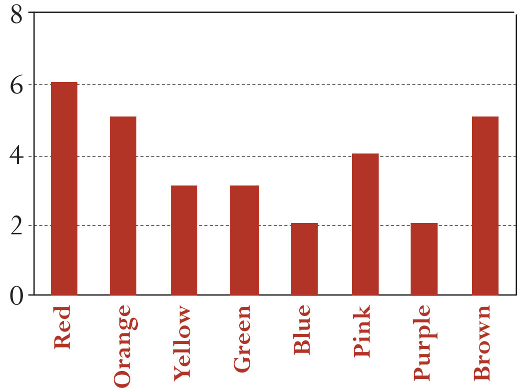
Nearly all students tried to respond to this question; across the OECD on average only 4% of students failed to respond.

Note that this unit includes one other question (INTERNET RELAY CHAT – Question 2) and this is presented in the section *Examples of difficult questions in PISA 2003 mathematics*.

COLOURED CANDIES

Question 1: COLOURED CANDIES

Robert's mother lets him pick one candy from a bag. He can't see the candies. The number of candies of each colour in the bag is shown in the following graph.



What is the probability that Robert will pick a red candy?

A 10%

B 20%

C 25%

D 50%

COLOURED CANDIES – Question 1 illustrates Level 4 in PISA 2003 mathematics and has a difficulty of 549 PISA score points. On average across the OECD countries, 50% of students were able to do this successfully.

Context: *Personal* – Many students can relate to this context through previous experiences, and such experiences involve some kind of probabilistic reasoning as young children do prefer certain colours or flavours. And they realise that certain colours are less abundant than others. Perhaps the problem lacks some authenticity for students at age 15, but the underlying concepts are valuable and relevant.

Content area: *Uncertainty* – this problem represents a wide array of problems that involve some thinking about chance. This problem measures an important aspect of *mathematical literacy* through its presentation of a more or less realistic situation that elicits probabilistic thinking, and its demand that students make direct and explicit connections between the context and a standard mathematical representation of a key aspect of the context – namely a bar chart representing the frequency distribution by colour of candies in the bag. This question formalises dealing with uncertainty in a fairly straightforward way.

Competency cluster: *Reproduction* – a complex and demanding combination of individual Reproduction competencies

The question requires students to:

- Identify relevant information from the graph (there are 6 red candies).
- Identify and calculate from the graph the total number of candies (6 + 5 + 3 + 3 + 2 + 4 + 2 + 5 or altogether 30 candies).
- Produce a basic probability calculation to get to the answer: 6 out of 30 is 20%.

Successful students answered "B" (20%). Students were most successful on this question in Iceland (76%), Korea (73%), the partner economy Hong Kong-China (72%), the Netherlands (69%) and Denmark (66%).

On average across the OECD countries, only 2% of students did not respond to this question.

LITTER

Question 1: LITTER

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

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

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

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

LITTER – Question 1 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 551 PISA score points. On average across OECD countries 52% of students were able to do this successfully.

Context: *Scientific*

Content area: *Uncertainty* – this question aims to test whether students are able to reason correctly about how to represent numbers (data) appropriately and their skills to effectively communicate this. As such, the scoring of this question is important. There are two possible correct answers, but students only need to give one of these: Either students make an argument based on the large differences in magnitude of the numbers involved, and the resulting difficulty in displaying these; or students make an argument based on the variability of the data within the different categories, and the resulting uncertainty in constructing a display.

Competency cluster: *Reflection* – Visualising an argument or data in an appropriate, meaningful and convincing way, and conversely, judging such representations on their qualities are key aspects of mathematical literacy. This requires some kind of reflection on the available data.

The question requires students to:

- Interpret the data.
- Reflect on the data.
- Communicate the results of their reflection.

Successful students gave answers focusing on either the big variance in data or the variability of the data for some categories. The question places an emphasis on communication of results. An examination of student responses to this question illustrates this point. The following two student responses were scored as correct answers:

- "You will get a mess, one starts at 0.5 years and another one at more than hundred years."
- "You have to make a vertical axis that goes minimally to 100 years with small steps because you need to able to read "a couple of days"."

Other correct answers could include:

- "The length of the bar for "polystyrene cups" is undetermined."
- "You cannot make one bar for 1–3 years or one bar for 20–25 years."
- "The difference in the lengths of the bars of the bar graph would be too big."
- "If you make a bar with length 10 centimetres for polystyrene, the one for cardboard boxes would be 0.05 centimetres."

Students were most successful on this question in Korea (74% correct), Finland (73%), Iceland (71%), the partner economy Hong Kong-China (68%) and Norway (67%). On average across OECD countries 32% of students attempted to answer the question but gave an incorrect answer. This varied from less than 20% of students in Korea and Poland to 76% of students in the United States. Examples of incorrect answers include:

- "Because it will not work."
- "A pictogram is better."
- "You cannot verify the info."
- "Because the numbers in the table are only approximations."

Across the OECD countries on average, 16% of students did not respond to this question.

SCIENCE TESTS

Question 1: SCIENCE TESTS

In Mei Lin's school, her science teacher gives tests that are marked out of 100. Mei Lin has an average of 60 marks on her first four Science tests. On the fifth test she got 80 marks.

What is the average of Mei Lin's marks in Science after all five tests?

Average: ……………………

SCIENCE TESTS – Question 1 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 556 PISA score points. On average across the OECD countries, 47% of students were able to do this correctly.

Context: *Educational and occupational* – this is a very familiar context for many students.

Content area: *Uncertainty* – weighted average. The problem-solving process could be as follows: add the score of 80 marks to the existing average for the first four science tests, that is, 60 marks. So: 60 + 60 + 60 + 60 + 80 = 320. (Or: 4 x 60 plus 80). Then divide this number by 5 to get the answer of 64 marks.

By far the most common incorrect response to this question was the answer 70. It is clear that this answer is incorrect, and it seems plausible to assume that these students have not read the stem of the problem accurately enough and rushed to the conclusion that the requested answer was the simple average of 60 and 80 (calculated as 60 + 80 divided by 2) rather than a weighted average that recognises that the total of the first four test scores must be 240.

Competency cluster: *Reproduction* – the concept of average is tested by giving a problem with a very familiar context, with simple numbers. The fact that the average of the first four scores were given might have added to the complexity, as the frequency of the incorrect answer 70 suggests.

The question requires students to:

- Read carefully.
- Have a proper understanding of the mathematical concept of the "average".
- "Reverse engineer" the rule for calculating an average to find the new average. This involves both the mathematisation of the concept of average and mathematical manipulation of the result.

Successful students answered "64". The most successful students on this question were in the partner economy Hong Kong-China (75% correct), the partner economy Macao-China (69%), Korea (67%), Japan (63%) and Canada (60%).

Across the OECD countries on average 16% of students did not respond to this question.

EARTHQUAKE

Question 1: EARTHQUAKE

A documentary was broadcast about earthquakes and how often earthquakes occur. It included a discussion about the predictability of earthquakes.

A geologist stated: "In the next twenty years, the chance that an earthquake will occur in Zed City is two out of three."

Which of the following best reflects the meaning of the geologist's statement?

A $\frac{2}{3} \times 20 = 13.3$, so between 13 and 14 years from now there will be an earthquake in Zed City.

B $\frac{2}{3}$ is more than $\frac{1}{2}$, so you can be sure there will be an earthquake in Zed City at some time during the next 20 years.

C The likelihood that there will be an earthquake in Zed City at some time during the next 20 years is higher than the likelihood of no earthquake.

D You cannot tell what will happen, because nobody can be sure when an earthquake will occur.

EARTHQUAKE – Question 1 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 557 PISA score points. On average across OECD countries, 46% if students were able to do this successfully.

Context: *Scientific*

Content area: *Uncertainty* – statistical forecasting/predictions. This question illustrates an important part of *mathematical literacy*. Experts often make predictions, although these are seldom transparent or explicit. For example, expressions used to forecast the weather, such as "there is a 20% chance of rain tomorrow". The viewer or reader thinks that there is a good chance it will remain dry tomorrow, but cannot complain if it rains for a substantial part of the day. Intelligent and mathematically literate citizens should be able to reflect in a critical way on what is actually meant by such a prediction.

Competency cluster: *Reflection* – students need to consider a given statement and reflect upon the meaning of that statement and four possible responses. Ideally such a question would require students to explain the result of their reflection in their own words, but such answers would probably be difficult to score objectively. Therefore the format of multiple-choice has been chosen. This means an extra step for the students: they may reflect first, and try to connect the result of this process to one of the four possible responses. Alternatively, students may consider the four possible responses and try to judge which one is the most likely. In this case the Reflection process takes a slightly different form.

Successful students answered "C". Students were most successful on this question in Japan (68%), Korea (64%) and Finland and New Zealand (59%). It is interesting that quite a large number of students chose the wrong answer "D" – 22% on average across OECD countries and as many as 39% of students in the Slovak Republic and the partner country Serbia and 36% in the Czech Republic. This statement is arguably correct but is not an answer to the question asked.

On average across OECD countries 9% of students did not respond to this question.

CHOICES

Question 1: CHOICES

In a pizza restaurant, you can get a basic pizza with two toppings: cheese and tomato. You can also make up your own pizza with **extra** *toppings. You can choose from four different extra toppings: olives, ham, mushrooms and salami.*

Ross wants to order a pizza with two different **extra** *toppings.*

How many different combinations can Ross choose from?

Answer: .... combinations.

CHOICES – Question 1 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 559 PISA score points. On average across OECD countries, 49% of students were able to do this successfully.

Context: *Educational and occupational* – while the problem is located in an occupational setting, such a question is likely only to be found in a school mathematics classroom. Nevertheless the thinking involved is demanded in many situations and is clearly part of *mathematical literacy*.

Content area: *Quantity* – this problem belongs clearly to the field in mathematics called combinatorics. However, it is not necessary to use knowledge other than structured reasoning. From a mathematical point of view the problem is not too complex. However, from a reading point of view, it is. There is one pizza, with two basic ingredients, and four extra choices of which the student can choose two. A more or less structured and safe solution is to draw a basic pizza (represented here by B), and subsequently draw this pizza with all possibilities having one extra ingredient (and using the numbers 1 to 4 to represent the extras): B1, B2, B3 and B4. It is now possible to add the second extra ingredient that has to be different: (B11), B12, B13, B14, (B21), (B22), B23, B24, (B31), (B32), (B33), B34, (B41), (B42), (B43) and (B44). So, it is only possible to create 6 different pizzas: B12, B13, B14, B23, B24 and B34.

Competency cluster: *Connections* – students have clearly to mathematise the problem in the sense that they really have to read the text very precisely and identify the relevant information in a structured way. Next they have to come up with an answer that requires an organised and systematic way of thinking, making clear that all combinations have been found.

The question requires students to:

- Read and interpret a rather complex text.
- Identify the relevant mathematics.
- Develop a structured strategy to ensure finding all the answers.

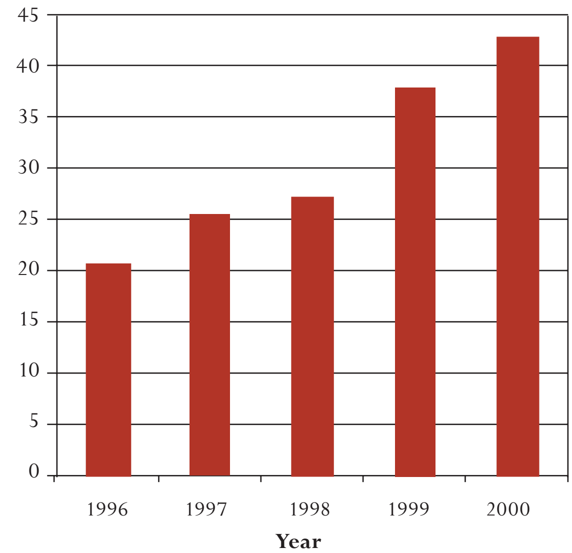
Successful students answered “6”. Students were most successful on this question in Japan (66%), Finland (60%), France and Korea (59%), the United Kingdom and Canada (58%).

On average across the OECD countries, most students attempted to answer this question – only 5% failed to respond.

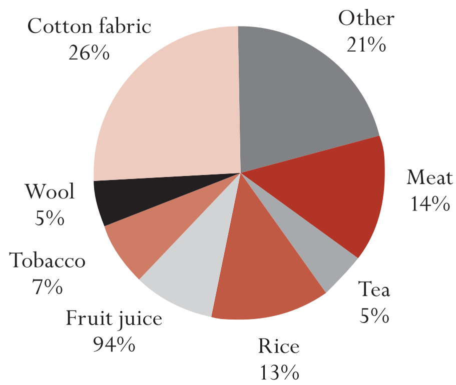
EXPORTS

The graphics below show information about exports from Zedland, a country that uses zeds as its currency.

Total annual exports from Zedland in millions of zeds, 1996-2000



Distribution of exports from Zedland in 2000



Question 2: EXPORTS

What was the value of fruit juice exported from Zedland in 2000?

A 1.8 million zeds.

B 2.3 million zeds.

C 2.4 million zeds.

D 3.4 million zeds.

E 3.8 million zeds.

EXPORTS – Question 2 illustrates Level 4 in PISA 2003 mathematics, with a difficulty of 565 PISA score points. On average across OECD countries, 48% of students were able to do this successfully.

Context: *Public* – The information society in which we live relies heavily on data, and data are often represented in graphics. The media use graphics often to illustrate articles and make points more convincingly. Reading and understanding this kind of information therefore is an essential component of *mathematical literacy*.

Content area: *Uncertainty* – focus on using data. The mathematical content consists of reading data from two graphs: a bar chart and a pie chart, comparing the characteristics of the two graphs, and combining data from the two graphs in order to be able to carry out a basic number operation resulting in a numerical answer.

Competency cluster: *Connections* – combine the information of the two graphics in a relevant way. This mathematisation process has some distinct phases. Students need to decode the different standard representations by looking at the total of annual exports of 2000 (42.6) and at the percentage of the Fruit Juice exports (9%) of this total. Students then need to connect these numbers by an appropriate numerical operation (9% of 42.6).

The question requires students to:

- Use mathematical insight to connect and combine two graphical representations.
- Apply the appropriate basic mathematical routine in the relevant way.

Successful students chose answer "E" (3.8 million zeds). Students were most successful on this question in the partner economy Hong Kong-China (69%), the partner economy Macao-China (63%), the Netherlands (62%) and Belgium and the Czech Republic (60%). The most common incorrect answer chosen by students was "C" (16% on average across OECD countries), followed by "A" (11%) and "B" (10%). See Chapter 6 for additional discussion.

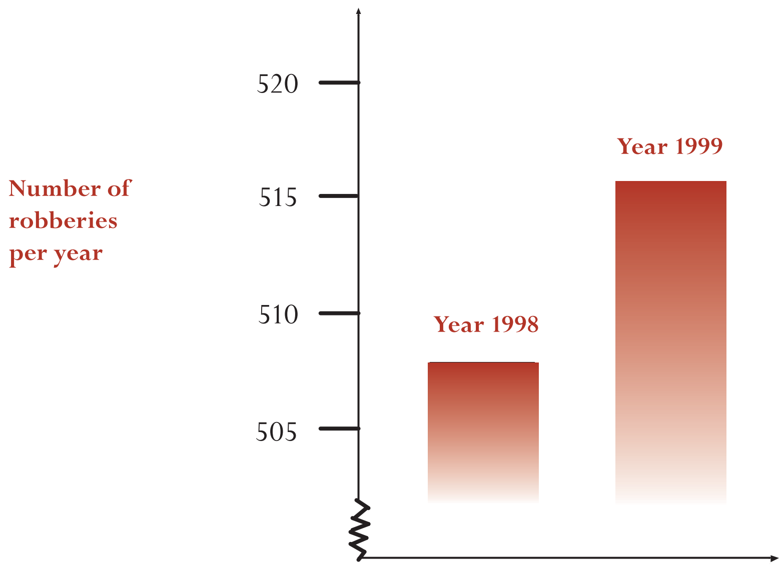
On average across the OECD countries 7% of students did not attempt to respond to this question, but this was the case for 16% of students in Italy and 20-21% of students in the partner countries Serbia and Uruguay.

ROBBERIES

Question 1: ROBBERIES

A TV reporter showed this graph and said:

"The graph shows that there is a huge increase in the number of robberies from 1998 to 1999."



Do you consider the reporter's statement to be a reasonable interpretation of the graph?

Give an explanation to support your answer.

ROBBERIES – Question 1 illustrates two levels of proficiency in PISA 2003 mathematics depending on whether students give partially or fully correct answers. Fully correct answers for this question illustrate Level 6, with a difficulty of 694 PISA score points. Here, a partially correct answer scored at 1 point illustrates performance at Level 4, with a difficulty of 577 PISA score points. On average across OECD countries, 28% of students were only capable of reaching this level of performance on ROBBERIES Q1.

Context: *Public* – The graph presented in this question was derived from a "real" graph with a similarly misleading message. The graph seems to indicate, as the TV reporter said: "a huge increase in the number of robberies". The students are asked if the statement fits the data It is very important to "look through" data and graphs as they are frequently presented in the media in order to function well in the knowledge society. This constitutes an essential skill in mathematical literacy. (See also the PISA Assessment Framework 2003, p. 105). Quite often designers of graphics use their skills (or lack thereof) to let the data support a pre-determined message, often with a political context. This is an example.

Content area: *Uncertainty* – analysis of a graph and interpretation of data. Understanding the issues related to misinterpretation of data. (In this graph the inappropriate cut in the y-axis indicates quite a large

increase in the number of robberies, but the absolute difference between the number of robberies in 1998 and 1999 is far from dramatic).

Competency cluster: *Connections* – reasoning and interpretation competencies, together with communication skills.

The question requires students to:

- Understand and decode a graphical representation in a critical way.
- Make judgments and find appropriate argumentation based on mathematical thinking and reasoning (interpretation of data).
- Use some proportional reasoning in a statistical context and a non-familiar situation.
- Communicate effectively their reasoning process.

Students were considered partially correct when they indicated that the statement is not reasonable, but fail to explain their judgment in appropriate detail. Their reasoning only focuses on an increase given by an exact number of robberies in absolute terms, but not in relative terms. In some cases, students may communicate their answers ineffectively leaving their answers open to interpretation. For example: "an increase of around 10 is not large" could mean something different to "an increase from 508 to 515 is not large". The second answer shows the actual numbers, and thus could indicate that the increase is small due to the large numbers involved, but the first answer does not show this line of reasoning. Examples of partially correct answers include:

- Not reasonable. It increased by about 10 robberies. The word "huge" does not explain the reality of the increased number of robberies. The increase was only about 10 and I wouldn't call that "huge".
- From 508 to 515 is not a large increase.
- No, because 8 or 9 is not a large amount.
- Sort of. From 507 to 515 is an increase, but not huge.

Very few students in each country answered that the interpretation was not reasonable, but made an error in calculating the percentage increase. Such answers were also considered to be partially correct.

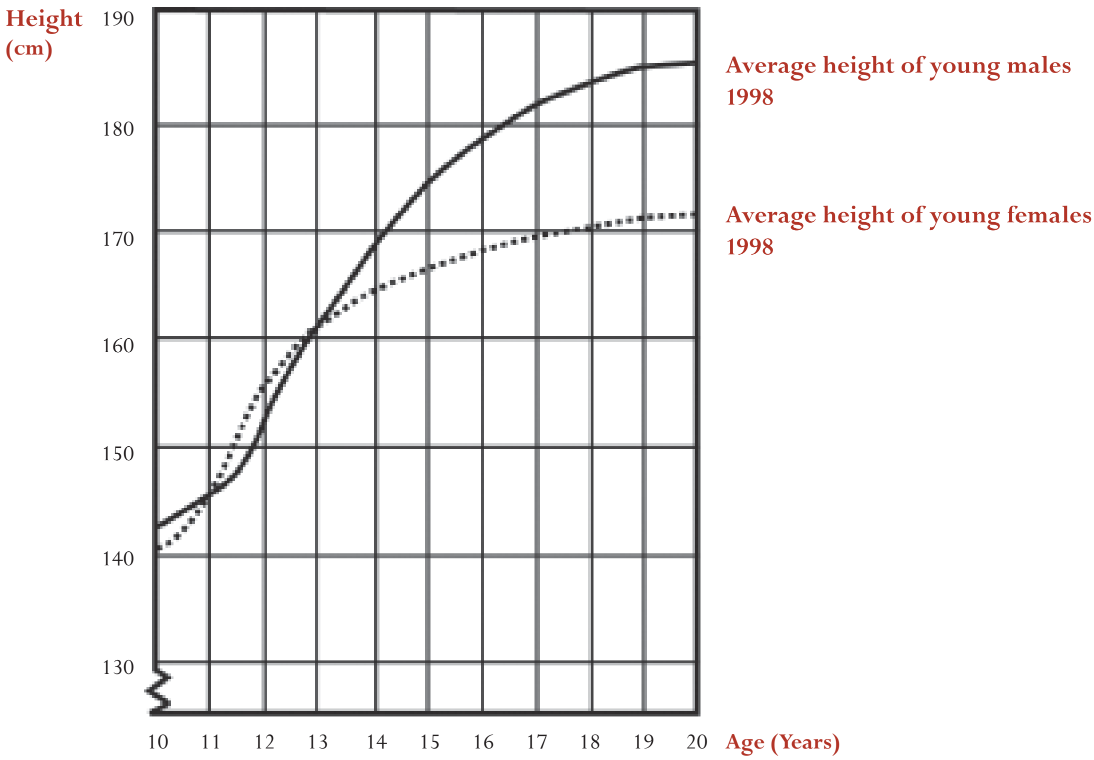
The following countries have the largest proportions of students who gave partially correct answers to this question: Finland (38%), Canada and Ireland (37%), the United Kingdom and Australia (36%) and Japan (35%).

Across the OECD countries on average, 15% of students did not respond to this question. This was the case for 30% of students in Greece, 28% in the Slovak Republic and 20% in Turkey, Mexico and Luxembourg, and for between 26 and 35% in the partner countries Serbia, Brazil, Uruguay, the Russian Federation, Tunisia and Indonesia.

GROWING UP

Youth grows taller

In 1998 the average height of both young males and young females in the Netherlands is represented in this graph



Question 3: GROWING UP

Explain how the graph shows that on average the growth rate for girls slows down after 12 years of age.

....

....

....

GROWING UP – Question 3 illustrates Level 4 in PISA 2003 mathematics with a difficulty of 574 PISA score points. On average across OECD countries, 45% of students were able to do this successfully.

Context: *Scientific*

Content area: *Change and relationships* – focus on the relationship between age and height. The mathematical concept of “decreasing growth”. This is used often in the media, but seldom properly understood. The problem is the combination of “growing” and “slowing down”, following the language used in the question.

In mathematical terms: the graphs should become less “steep”. Even more mathematically: the slope (or gradient) would decrease.

Competency cluster: *Connections* – solve a problem in a non-routine situation, although still involving familiar settings. Students need to think and reason (what does the question mean in mathematical terms?), make an argument, and communicate this in a proper way (which is not trivial here). Students also need to solve the problem and decode the graph. The question is definitely not familiar and demands the intelligent linking of different ideas and information.

The question requires students to:

- Show mathematical insight.
- Analyse different growth curves.
- Evaluate the characteristics of a data set, represented in a graph.
- Note and interpret the different slopes at various points of the graphs.
- Reason and communicate the results of this process, within the explicit models of growth.

Successful students were able to read the graph correctly to determine that growth starts to diminish at age 12, or a bit before that age, and communicate this observation. Students were most successful on this question in the Netherlands (78%), Finland (68%) and Canada and Belgium (64%), where at least 88% of students responded to the question, compared to 79% on average across OECD countries. However, in some OECD countries significant proportions of students did not attempt to respond to this question, notably in Austria (44%) and Greece (43%).

In all countries successful students gave answers ranging from daily-life language to more mathematical language involving the reduced steepness, or they compared the actual growth in centimetres per year. Among the OECD countries, the most common correct answers were given in daily-life language. For example:

- It no longer goes straight up, it straightens out.
- The curve levels off.
- It is flatter after 12.
- The girls’ line starts to even out and the boys’ line just gets bigger.
- It straightens out and the boys’ graph keeps rising.

This was the case for at least 70% of correct answers in 24 of the OECD countries, but only 39% in Korea and 49% in Austria. In Korea 56% of the correct answers were communicated in mathematical language, where students used terms such as “gradient”, “slope”, or “rate of change”. This was the case for between 21 and 26% of correct answers in New Zealand, Turkey, Hungary, Canada, Japan and the Slovak Republic. In Austria, 34% of correct answers consisted of students comparing the actual growth. Examples of such answers include:

- From 10 to 12 the growth is about 15 cm, but from 12 to 20 the growth is only about 17 cm.
- The average growth rate from 10 to 12 is about 7.5 cm per year, but about 2 cm per year from 12 to 20 years.

Such answers comparing the actual growth also comprised a significant proportion of the correct answers in the following OECD countries: Mexico (26%), Greece (23%), France and Turkey (19%).

The most common error that students made was to give an answer that did not refer to the graph, for example “girls don’t grow much after 12”. However, around 40% of the incorrect answers given in France, Korea and Poland did refer to the graph, but simply indicated that the female height drops below the male height, without referring to the steepness of the female gradient.

EXCHANGE RATE

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

Question 3: EXCHANGE RATE

During these 3 months the exchange rate had changed from 4.2 to 4.0 ZAR per SGD.

Was it in Mei-Ling's favour that the exchange rate now was 4.0 ZAR instead of 4.2 ZAR, when she changed her South African rand back to Singapore dollars? Give an explanation to support your answer.

EXCHANGE RATE – Question 3 illustrates Level 4 in PISA 2003 mathematics with a difficulty of 586 PISA score points. On average across OECD countries, 40% of students were able to do this successfully.

Context: *Public* – currency exchange associated with international travel

Content area: *Quantity* – quantitative relationships with money, procedural knowledge of number operations (multiplication and division)

Competency cluster: *Reflection* – students have to reflect on the concept of exchange rate and its consequences in this particular situation. This question illustrates the process of mathematisation. First students need to identify the relevant mathematics involved in this real-world problem. Although all the required information is explicitly presented in the question this is a somewhat complex task. Reducing the information in the question to a problem within the mathematical world places significant demands on students. Students need to think and reason flexibly (how do we find?), form an argument (how are the objects related?) and solve the mathematical problem. Combining these three competencies requires students to reflect on the process needed to solve the problem. Finally students need to communicate a real solution and explain the conclusion.

The question requires students to:

- Interpret a non-routine mathematical relationship (a specified change in the exchange rate for 1 Singapore Dollar/1 South African Rand).
- Reflect on this change.
- Use flexible reasoning to solve the problem.
- Apply some basic computational skills or quantitative comparison skills.
- Construct an explanation of their conclusion.

Students were most successful on this question in the partner country Liechtenstein (64%), Canada (58%), Belgium (55%), the partner economies Macao-China and Hong Kong-China (53%), Sweden, Finland and France (51%). Less than 20% of students answered this question correctly in Mexico and Turkey and in the partner countries Indonesia, Brazil, Thailand and Tunisia.

Forty-two percent of students responded to this question but gave a wrong answer, on average across OECD countries. In some cases, students answered “yes” but failed to give an adequate explanation or gave no explanation at all. For example:

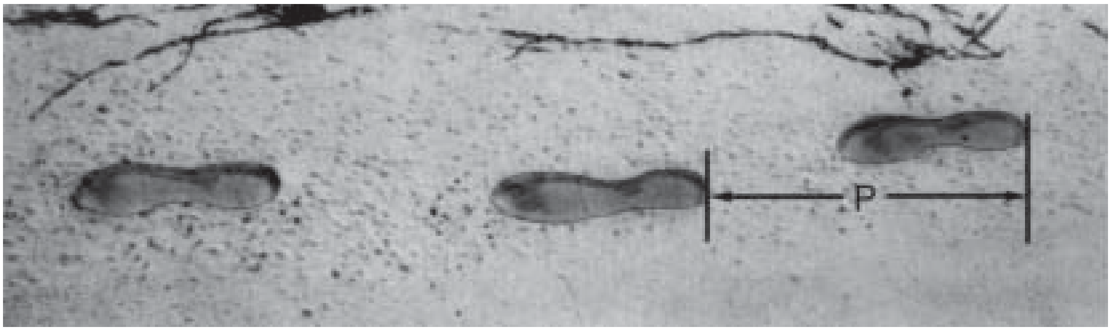
- Yes, a lower exchange rate is better.
- Yes it was in Mei-Ling’s favour, because if the ZAR goes down, then she will have more money to exchange into SGD.
- Yes it was in Mei-Ling’s favour.

This was the case for 54% of the wrong answers in France and between 40% and 49% of the wrong answers in Ireland, New Zealand, Portugal, Switzerland, Australia, Austria, Greece, Finland, Spain, Luxembourg, Japan and the partner countries/economies the Russian Federation and Hong Kong-China.

A further 17% did not respond to the question and this was between 27 and 29% in Mexico, Italy, Portugal, Turkey, Greece, and the partner country Serbia, 35% in the partner country Tunisia and 42% in the partner country Brazil.

Note that this unit includes two other questions (EXCHANGE RATE – Question 1 and EXCHANGE RATE – Question 2) and these are presented in the section *Examples of easy questions in PISA 2003 mathematics*.

WALKING



The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

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

n = *number of steps per minute, and*

P = *pacelength in metres.*

Question 3: WALKING

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

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

WALKING – Question 3 illustrates three levels of proficiency in PISA 2003 mathematics depending on whether students give partially or fully correct answers. Fully correct answers for this item illustrate the high part of Level 6, with a difficulty of 723 PISA score points. There are two levels of partially correct answers: the higher level illustrates the higher part of Level 5, with a difficulty of 666 PISA score points (just 3 points below the boundary with Level 6) and the lower level illustrates the top part of Level 4, with a difficulty of 605 PISA score points (just 2 points below the boundary with Level 5). On average across OECD countries, 21% of students were able to do this successfully. Here we discuss the lower level of credit for WALKING Q3. This lower level received a score of 1 point and accounted for the 21% of responding students who were unable to do anything more on this problem.

Context: *Personal*

Content area: *Change and relationships* – the relationship between the number of steps per minute and pace-length. Conversion of measurement from m/min to km/hr.

The mathematical routine needed to solve the problem successfully is substitution in a simple formula (algebra), and carrying out a non-routine calculation. The first step in the solution process requires students to calculate the number of steps per minute when the pace-length is given (0.8 m). This requires proper substitution: $n/0.80 = 140$ and the observation that this equals: $n = 140 \times 0.80$ which in turn is 112 (steps per minute). The problem requires more than just routine operations: first substitution in an algebraic expression, followed by manipulating the resulting formula, in order to be able to carry out the required calculation. The next step is to go beyond the observation that the number of steps is 112. The question asks for the speed in m/minute: per minute he walks $112 \times 0.80 = 89.6$ meters; so his speed is 89.6 m/minute. The final step is to transform this speed from m/minute into km/h, which is a more commonly used unit of speed. This involves relationships among units for conversions within systems of units and for rates which is part of the measurement domain. Solving the problem also requires decoding and interpreting basic symbolic language in a less known situation, and handling expressions containing symbols and formulae.

Competency cluster: *Connections* – The problem is rather complex in the sense that not only is use of a formal algebraic expression required, but also doing a sequence of different but connected calculations that need proper understanding of transforming formulas and units of measures.

The question requires students to:

- Complete the conversions.
- Provide a correct answer in both of the requested units.

Students scoring at the lower level of partially correct answers includes those who wrote an expression that showed they had understood the formula and correctly substituted the appropriate values into it, finding the number of steps per minute. Such answers include:

- $n = 140 \times .80 = 112$. No further working out is shown or incorrect working out from this point.
- $n = 112$, 0.112 km/h.
- $n = 112$, 1120 km/h.
- 112 m/min, 504 km/h.

On average across OECD countries 20% of students were only able to achieve this lower level of partially correct answer. This was the case for 35% of students in the United States, 33% of students in Canada, 31% of students in the Slovak Republic and 30% of students in Greece.

EXAMPLES OF DIFFICULT MATHEMATICS QUESTIONS FROM PISA 2003

Twenty-seven mathematics questions from PISA 2003 lie in Levels 5 and 6 of the literacy scale. Nine of these relatively difficult questions have been released, and these are listed in Table 3.3 along with the difficulty of each question on the PISA mathematics scale, and other key framework characteristics. Note that two of these questions had one or more levels of partial credit associated with them, and the full and partial credit score points are listed in the table accordingly.

The absence of the *reproduction* competency cluster amongst the more difficult released questions seen in Table 3.3 is typical given the generally greater cognitive demands imposed by questions at Levels 5 and 6 of the PISA mathematics scale. All but one of the most difficult released questions are classified in either the *connections* or *reflection* competency clusters. The exception is a question requiring the application of routine knowledge and procedures, but using algebra in a real-world context. This is what makes the question more difficult than might otherwise be expected for questions in the *reproduction* competency cluster. The need to reflect substantively on the situation presented or on the solution obtained is a key challenge that tends to immediately make test questions more difficult than those for which such a demand is not made. The need to make connections among problem elements in order to solve a problem also makes questions more difficult compared to questions requiring the simple reproduction of practised knowledge and questions limited to the direct treatment of unconnected pieces of information.

Questions from all content areas, and from each of the context categories appear among the most difficult PISA mathematics questions, however only one of these is in the *quantity* area, and that question is not among the released set. Other more difficult *quantity* questions were developed for possible inclusion, but were not selected for the final PISA 2003 mathematics assessment.

Further, five of the nine questions require students to provide an extended response (either an extended sequence of calculations or an explanation or written argument in support of the conclusion). The most difficult questions typically have two features: the response structure is left open for the student, and active communication is required.

Finally, before introducing the more difficult released questions, the role of reading demand should be noted. Over half of the 27 most difficult questions in the PISA 2003 mathematics assessment are classified as “long”, meaning they contain more than 100 words. The released questions listed in Table 3.3 also reflect this observation, since five of the nine questions are in this category. Reading demand is an important component of question difficulty.

Figures showing country level performances on many of these difficult items are found in Annex A1, Figures A.1.24 through A.1.31.

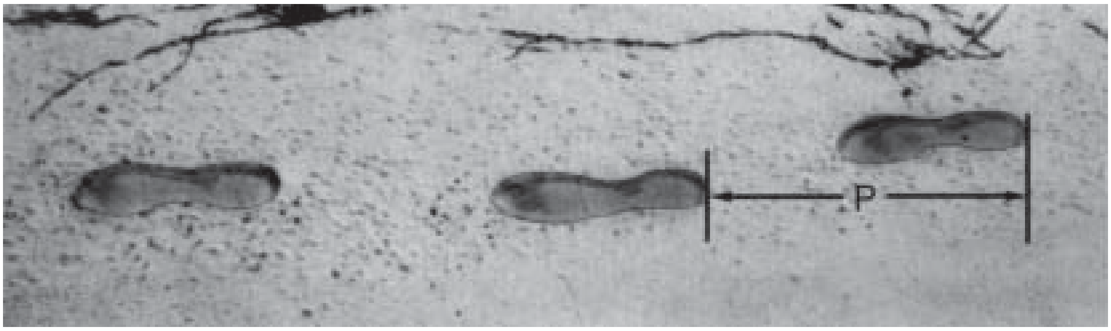
Table 3.3

Characteristics of the most difficult questions released from the PISA 2003 mathematics assessment

Item code	Question	OECD average percent correct	Location on PISA scale (PISA score points): Question	Location on PISA scale (PISA score points): Full/partial credit points	Traditional topic	Content area ("Overarching idea")	Competency cluster	Context ("Situation")	Length of question[1]	Response format
M124Q01	WALKING – Question 1	36	611	**611**	Algebra	Change and relationships	Reproduction	Personal	Medium	Extended Response
M702Q01	SUPPORT FOR THE PRESIDENT – Question 1	36	615	**615**	Data	Uncertainty	Connections	Public	Long	Extended Response
M513Q01	TEST SCORES – Question 1	32	620	**620**	Data	Uncertainty	Connections	Educational and occupational	Long	Extended Response
M710Q01	FORECAST OF RAIN – Question 1	34	620	**620**	Data	Uncertainty	Connections	Public	Long	Multiple Choice
M402Q02	INTERNET RELAY CHAT – Question 2	29	636	**636**	Measurement	Change and relationships	Reflection	Personal	Long	Short Answer
M704Q02	THE BEST CAR – Question 2	25	657	**657**	Algebra	Change and Relationships	Reflection	Public	Long	Short Answer
M124Q03	WALKING – Question 3	21	665	605 **666** **723**	Algebra	Change and relationships	Connections	Personal	Medium	Extended Response
M266Q01	CARPENTER – Question 1	20	687	**687**	Measurement	Space and shape	Connections	Educational and occupational	Medium	Complex Multiple Choice
M179Q01	ROBBERIES – Question 1[2]	30	635	577 **694**	Data	*Uncertainty*	*Connections*	*Public*	Short	Extended Response

1. Short questions contain fewer than 50 words. Medium-length questions contain 51 to 100 words. Long questions contain more than 100 words. Length of question in relation to question difficulty is discussed in detail in Chapter 5.

WALKING



The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

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

n = number of steps per minute, and

P = pacelength in metres.

Question 1: WALKING

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

Show your work.

WALKING – Question 1 illustrates Level 5 in PISA 2003 mathematics, with a difficulty of 611 PISA score points (just four points beyond the boundary with Level 4). On average across OECD countries, 36% of students were able to do this successfully.

Context: *Personal* – Everyone has seen his or her own footsteps printed in the ground (whether in sand or mud) at some moment in life, most likely without realising the kind of relations that exist in the way these patterns are formed (although many students will have an intuitive feeling that if the pace-length increases, the number of steps per minute will decrease). To reflect on and realise the embedded mathematics in such daily phenomena is part of acquiring *mathematical literacy*.

Content area: *Change and relationships* – the relationship between the number of steps per minute and the pace-length. This relationship was derived from observing many different people walking steadily at their natural pace in a variety of situations. The mathematical content could be described as belonging clearly to algebra. Students need to solve the problem successfully by substituting in a simple formula and carrying

out a routine calculation ($70/p = 140$) to find the value of p. The students need to carry out the actual calculation in order to get full credit.

Competency cluster: *Reproduction* – The competencies needed involve reproduction of practised knowledge, the performance of routine procedures, application of standard technical skills, manipulation of expressions containing symbols and formulae in standard form, and carrying out computations.

The question requires students to:

- Use a formal algebraic expression to solve a problem.

Successful students gave both the formula and the correct result. Examples of correct answers are:

0.5 m or 50 cm, ½ (unit not required).

$70/p = 140$

$70 = 140p$

$p = 0.5$

$70/140$

Students were most successful on this question in the partner economy Hong Kong-China (62%), the partner economy Macao-China (60%), the partner country the Russian Federation (54%), the Netherlands (52%) and the Slovak Republic (52%). On average across OECD countries 22% of students gave the correct formula but did not give the correct answer. This was the case for 48% of students in the United States, 35% of students in Ireland, 32% of students in Portugal and Luxembourg, 31% of students in Iceland, Poland and the partner country Indonesia. The original intention was to consider answers to be partially correct if students just gave the formula, but not the result or an incorrect result. However, the average ability of these students was not sufficiently higher than that of students who simply gave an incorrect answer. So no credit was awarded to students who only gave the formula.

Question 3: WALKING

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

Calculate Bernard's walking speed in metres per minute and in kilometres per hour.

Show your work.

WALKING – Question 3 illustrates three levels of proficiency in PISA 2003 mathematics depending on whether students give partially or fully correct answers. Fully correct answers for this item illustrate the high part of Level 6, with a difficulty of 723 PISA score points. There are two levels of partially correct answers: the higher level illustrates the higher part of Level 5, with a difficulty of 666 PISA score points (just three points short of the boundary with Level 6) and the lower level discussed in the previous section dealing with Level 4, with a difficulty of 605 PISA score points. On average across OECD countries, 21% of students were able to solve this problem successfully.

Context: *Personal*

Content area: *Change and relationships* – the relationship between the number of steps per minute and pace-length. Conversion of measurement from m/min to km/hr.

The mathematical routine needed to solve the problem successfully is substitution in a simple formula (algebra), and carrying out a non-routine calculation. The first step in the solution process requires students to calculate the number of steps per minute when the pace-length is given (0.8 m). This requires proper substitution: n/0.80 = 140 and the observation that this equals: $n = 140 \times 0.80$ which in turn is 112 (steps per minute). The problem requires more than just routine operations: first substitution in an algebraic expression, followed by manipulating the resulting formula, in order to be able to carry out the required calculation. The next step is to go beyond the observation that the number of steps is 112. The question asks for the speed in m/minute: per minute he walks 112 × 0.80 = 89.6 meters; so his speed is 89.6 m/minute. The final step is to transform this speed from m/minute into km/h, which is a more commonly used unit of speed. This involves relationships among units for conversions within systems of units and for rates which is part of the measurement domain. Solving the problem also requires decoding and interpreting basic symbolic language in a less known situation, and handling expressions containing symbols and formulae.

Competency cluster: *Connections* – The problem is rather complex in the sense that not only is use of a formal algebraic expression required, but also doing a sequence of different but connected calculations that need proper understanding of transforming formulas and units of measures.

The question requires students to:

- Complete the conversions.
- Provide a correct answer in both of the requested units.

Successful students gave correct answers for both metres/minute and km/hour (although the units were not required). An example of a fully correct answer:

> $n = 140 \times .80 = 112$.
>
> Per minute he walks 112 × .80 metres = 89.6 metres.
>
> His speed is 89.6 metres per minute.
>
> So his speed is 5.38 or 5.4 km/hr.

Also fully correct answers do not need to show the working out (*e.g.* 89.6 and 5.4) and errors due to rounding are acceptable (*e.g.* 90 metres per minute and 5.3 km/hr [89 × 60]). However an answer of 5376 must specify the unit (m/hour) to be considered fully correct.

Students were most successful on this question in the partner economy Hong Kong-China (19% correct), Japan (18%), Belgium (16%), the Netherlands, the partner countries/economies Macao-China and Liechtenstein (15%) and Finland and Switzerland (14%).

Some students were able to find the number of steps per minute and make some progress towards converting this into the more standard units of speed asked for. However, their answers were not entirely complete or fully correct. Examples of partially correct answers include where:

- Students fail to multiply by 0.80 to convert from steps per minute to metres per minute. For example, his speed is 112 metres per minute and 6.72 km/hr.
- Students give the correct speed in metres per minute (89.6 metres per minute) but conversion to kilometres per hour is incorrect or missing.
- Students explicitly show the correct method, but make minor calculation error(s). No answers correct.

  n = 140 x .8 = 1120; 1120 × 0.8 = 896. He walks 896 m/min, 53.76 km/h.

  n = 140 x .8 = 116; 116 × 0.8 = 92.8. 92.8 m/min, 5.57 km/h.

- Students only give 5.4 km/hr, but not 89.6 metres/minute (intermediate calculations not shown).

On average across OECD countries 9% of students gave one of the above answers and were awarded the higher level of partially correct answers. This was the case for 30% of students in the partner economy Hong Kong-China, 26% in the partner economy Macao-China and 20% in Japan. Among these students the most common error was not to convert the number of steps into metres (this concerned around 70% of the higher level partially correct answers in Japan and the partner economies Hong Kong-China and Macao-China). Indeed, failure to convert the number of steps into metres was the most common reason why students just fell short of fully correct answers in 19 of the OECD countries. Conversely, the majority of students with higher level partially correct answers in Hungary, the Slovak Republic, Greece and Italy failed to convert from metres per minute into kilometres per hour (this concerned around 60% of such answers).

SUPPORT FOR THE PRESIDENT

Question 1: SUPPORT FOR THE PRESIDENT

In Zedland, opinion polls were conducted to find out the level of support for the President in the forthcoming election. Four newspaper publishers did separate nationwide polls. The results for the four newspaper polls are shown below:

Newspaper 1: 36.5% (poll conducted on January 6, with a sample of 500 randomly selected citizens with voting rights)

Newspaper 2: 41.0% (poll conducted on January 20, with a sample of 500 randomly selected citizens with voting rights)

Newspaper 3: 39.0% (poll conducted on January 20, with a sample of 1000 randomly selected citizens with voting rights)

Newspaper 4: 44.5% (poll conducted on January 20, with 1000 readers phoning in to vote).

Which newspaper's result is likely to be the best for predicting the level of support for the President if the election is held on January 25? Give two reasons to support your answer.

SUPPORT FOR THE PRESIDENT – Question 1 illustrates Level 5 in PISA 2003 mathematics, with a difficulty of 615 PISA score points. On average across OECD countries, 36% of students were able to do this successfully.

Context: *Public* – This problem illustrates an important aspect of mathematical literacy: the ability for citizens to critically judge presentations with a mathematical background. This is especially important for presentations like opinion polls that seem to be used increasingly in this media-centred society. Particularly when the articles or television items mention that the prediction or poll may not be "representative", or is not taken randomly, or is not "fair" in any other way. It is important not to simply accept such statements and results without looking closely at the data in the context of how they were collected.

Content area: *Uncertainty* – sampling. There are four important characteristics to evaluate the samples in the question: the more recent survey tends to be better, the survey should be taken from a large sample, it should be a random sample, and of course only respondents who are eligible to vote should be considered. To gain full credit students needed to come up with two of these four arguments, and therefore choose Newspaper 3.

Competency cluster: *Connections* – although some reflection may be helpful to the students. As well as needing a good understanding of sampling, students need to read a rather complex text and understand each of the four possibilities.

The question requires students to:

- Understand the text.
- Understand conceptually different aspects of sampling.
- Produce and write the reasons for choosing the answer given.
- Successful students answered "Newspaper 3" and gave at least two valid reasons to justify this conclusion. Possible reasons include: The poll is more recent, with larger sample size; a random selection of the sample; and only voters were asked. If students gave additional information (including irrelevant or

incorrect information) this was simply ignored. The essential was that they had given the correct answer with two valid reasons. Examples of correct answers include:

- Newspaper 3, because they have selected more citizens randomly with voting rights.
- Newspaper 3 because it has asked 1000 people, randomly selected, and the date is closer to the election date so the voters have less time to change their mind.
- Newspaper 3 because they were randomly selected and they had voting rights.
- Newspaper 3 because it surveyed more people closer to the date.
- Newspaper 3 because the 1000 people were randomly selected.

Students were most successful on this question in the partner economy Hong Kong-China (48%), France and Japan (47%), Finland, Canada, Australia, the Netherlands and Korea (46%) and New Zealand (45%). On average across OECD countries, 7% of students answered “Newspaper 3”, but did not give an adequate explanation or gave no explanation. This was the case for less than 5% of students in Poland, Turkey, Japan and the partner economy Hong Kong-China.

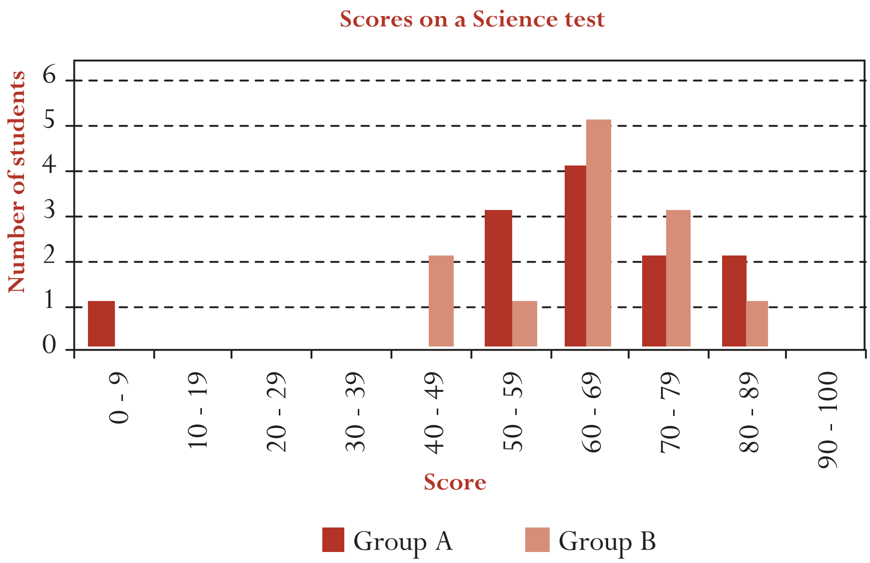
TEST SCORES

Question 1: TEST SCORES

The diagram below shows the results on a Science test for two groups, labelled as Group A and Group B.

The mean score for Group A is 62.0 and the mean for Group B is 64.5. Students pass this test when their score is 50 or above.

Looking at the diagram, the teacher claims that Group B did better than Group A in this test.



The students in Group A don't agree with their teacher. They try to convince the teacher that Group B may not necessarily have done better.

Give one mathematical argument, using the graph, that the students in Group A could use.

TEST SCORES – Question 1 illustrates Level 5 in PISA 2003 mathematics, with a difficulty of 620 PISA score points. On average across OECD countries 32% of students were able to do this successfully.

Context: *Educational and occupational* – the educational context of this item is one that all students are familiar with: comparing test scores. In this case a science test has been administered to two groups of students: A and B. The results are given to the students in two different ways: in words with some data embedded and by means of two graphs in one grid.

Content area: *Uncertainty* – the field of exploratory data analysis. Knowledge of this area of mathematics is essential in the information society in which we live, as data and graphical representations play a major role in the media and in other aspects of daily experiences.

Competency cluster: *Connections* – includes competencies that not only build on those required for the *reproduction* competency cluster (like encoding and interpretation of simple graphical representations) but also require reasoning and insight, and in particular, mathematical argument. The problem is to find

arguments that support the statement that Group A actually did better than group B, given the counter-argument of one teacher that group B did better – on the grounds of the higher mean for group B. Actually the students have a choice of at least three arguments here. The first one is that more students in group A pass the test; a second one is the distorting effect of the outlier in the results of group A; and finally Group A has more students that scored 80 or over. Another important competency needed is explaining matters that include relationships.

This question requires students to:

- Apply statistical knowledge in a problem situation that is somewhat structured and where the mathematical representation is partially apparent.
- Use reasoning and insight to interpret and analyse the given information.
- Communicate their reasons and arguments.

Many students did not respond to this question – 32% on average across OECD countries. Although this varies significantly among countries from 10% in the Netherlands and 13% in Canada to 49% in Mexico and the partner country Uruguay and 53% in Italy and 70% in the partner country Serbia.

Successful students gave one valid argument. Valid arguments could relate to the number of students passing, the disproportionate influence of the outlier, or the number of students with scores in the highest level. For example:

- More students in Group A than in Group B passed the test.
- If you ignore the weakest Group A student, the students in Group A do better than those in Group B.
- More Group A students than Group B students scored 80 or over.

Students were most successful on this question in the partner economies Hong Kong-China (64%) and Macao-China (55%) and in Japan (55%), Canada (47%), Korea (46%) and Belgium (44%).

On average across OECD countries, 33% of students responded to the question, but gave an incorrect answer. These included answers with no mathematical reasons, or wrong mathematical reasons, or answers that simply described differences but were not valid arguments that Group B may not have done better. For example:

- Group A students are normally better than Group B students in science. This test result is just a coincidence.
- Because the difference between the highest and lowest scores is smaller for Group B than for Group A.
- Group A has better score results in the 80-89 range and the 50-59 range.
- Group A has a larger inter-quartile range than Group B.

A significant proportion of students did not respond to this question (35% on average across OECD countries), although this varied from 11% in the Netherlands and 14% in Canada to 58% in Italy and Mexico, over 60% in the partner countries Tunisia and Uruguay and 73% in the partner country Serbia.

FORECAST OF RAINFALL

Question 1: FORECAST OF RAINFALL

On a particular day, the weather forecast predicts that from 12 noon to 6 pm the chance of rainfall is 30%. Which of the following statements is most likely to reflect the intended meaning of this forecast?

A 30% of the land in the forecast area will get rain.

B 30% of the 6 hours (a total of 108 minutes) will have rain.

C For the people in that area, 30 out of every 100 people will experience rain.

D If the same prediction was given for 100 days, then about 30 days out of the 100 days will have rain.

E The amount of rain will be 30% of a heavy rainfall (as measured by rainfall per unit time).

FORECAST OF RAINFALL – Question 1 illustrates a Level 5 item and has a difficulty level of 620 on the PISA score scale. On average across OECD countries, 34% of students were able to do this successfully.

Context: *Public* – forecast of rain is connected to media presentations and probability of events being held or cancelled

Content area: *Uncertainty* – involves the interpretation of factors and procedures associated with interpreting a statement involving probabilities

Competency cluster: *Connections* – students have to reflect on the concept of probability contained in a statement and use it to judge the validity of a number of statements

This question requires students to:

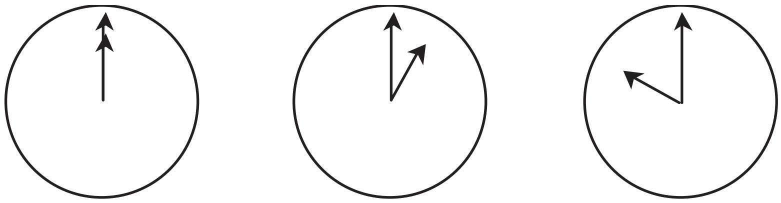
- Correctly interpret the given statement and connect it to the context described
- Use reflection and insight interpreting a standard probabilistic situation
- Compare and contrast the proposed communications based on information

Considerable variation was noted in student responses, ranging from a high of 54% correct in Korea and 49% correct in both Finland and partner country Liechtenstein to 11% in partner country Indonesia, 8% in Thailand, and 7% in partner country Tunisia.

INTERNET RELAY CHAT

Mark (from Sydney, Australia) and Hans (from Berlin, Germany) often communicate with each other using "chat" on the Internet. They have to log on to the Internet at the same time to be able to chat.

To find a suitable time to chat, Mark looked up a chart of world times and found the following:



Greenwich 12 Midnight **Berlin 1:00 AM** **Sydney 10:00 AM**

Question 2: INTERNET RELAY CHAT

Mark and Hans are not able to chat between 9:00 AM and 4:30 PM their local time, as they have to go to school. Also, from 11:00 PM till 7:00 AM their local time they won't be able to chat because they will be sleeping.

When would be a good time for Mark and Hans to chat? Write the local times in the table.

Place	Time
Sydney	
Berlin	

INTERNET RELAY CHAT – Question 2 illustrates Level 5 in PISA 2003 mathematics, with a difficulty of 636 PISA score points. On average across OECD countries, 29% were able to do this successfully.

Context: *Personal* – this assumes either that students are familiar at some level with chatting over the internet, and/or they know about time differences in this or another context. Given increasing globalisation and the enormous popularity of the internet this question really deals with *mathematical literacy*.

Content area: *Change and relationships* – time changes in different time zones.

Competency cluster: *Reflection* – rather high mathematisation skills are required to solve a non-routine problem. Students need to identify the relevant mathematics. Although the question seems rather straight-forward, and the numbers and the actual mathematical operations required are rather simple, the question is actually more complex. The students have to understand the way that time spent sleeping and at school constrains the times that could be suitable for communicating with each other. First students need to identify the times that could work for each of them separately. Then, students have to compare two "time-windows" to find a time that would work for both of them simultaneously. This involves performing the same time calculation as in Question 1 of this unit, but within a context constrained by the students'

analysis of the possibilities. It is worth noting that the question could have been made more complex had the problem been to identify the whole window of opportunity. But the question requests the student to find just one particular time that would work, giving the students the opportunity to use trial-and-error methods.

The question requires students to:

- Understand the question.
- Mathematise the question.
- Identify one time that will work.

Successful students gave an answer with any time (*e.g.* Sydney 17:00, Berlin 8:00) or interval of time satisfying the 9 hours time difference. These could be taken from one of the following intervals:

Sydney: 4:30 PM – 6:00 PM; Berlin: 7:30 AM – 9:00 AM

Sydney: 7:00 AM – 8:00 AM; Berlin: 10:00 PM – 11:00 PM

If students gave an interval of time this needed to satisfy the constraints in its entirety. Also, students who did not specify morning (AM) or evening (PM), but gave times that could otherwise be regarded as correct, were given the benefit of the doubt and their answers were considered correct. Between 36% and 42% of students were successful on this question in New Zealand, Australia, Switzerland, Ireland, Canada, the Netherlands and Belgium, as well as in the partner country Liechtenstein. On average across OECD countries, 52% of students gave an incorrect answer (e.g. only one correct time) and 19% of students did not respond to the question. The highest percentages of students not responding to the question were in Denmark (31%), Spain (30%) and the partner country Serbia (45%).

THE BEST CAR

A car magazine uses a rating system to evaluate new cars, and gives the award of "The Car of the Year" to the car with the highest total score. Five new cars are being evaluated, and their ratings are shown in the table.

Car	Safety Features (S)	Fuel Efficiency (F)	External Appearance (E)	Internal Fittings (T)
Ca	3	1	2	3
M2	2	2	2	2
Sp	3	1	3	2
N1	1	3	3	3
KK	3	2	3	2

The ratings are interpreted as follows:

3 points = Excellent

2 points = Good

1 point = Fair

To calculate the total score for a car, the car magazine uses the following rule, which is a weighted sum of the individual score points:

$$\text{Total Score} = (3 \times S) + F + E + T$$

Question 2: THE BEST CAR

The manufacturer of car "Ca" thought the rule for the total score was unfair.

Write down a rule for calculating the total score so that Car "Ca" will be the winner.

Your rule should include all four of the variables, and you should write down your rule by filling in positive numbers in the four spaces in the equation below.

Total score = ……… × S + ……… × F + ……… × E + ……… × T

THE BEST CAR – Question 2 illustrates Level 5 in PISA 2003 mathematics with a difficulty of 657 PISA score points. On average across OECD countries, 25% of students were able to do this successfully.

Context: *Public* – an article in a car magazine is a very familiar context, especially for males. The underlying mathematics is relevant for males and females as everyone is presented with this kind of problem, that is, the evaluation of a consumer good using a rating system, whether it be cars, washing machines, coffee makers, etc. This is therefore an important part of *mathematical literacy*.

Content area: *Change and relationships* – the focus is on the relationship of numbers in a formula.

Competency cluster: *Reflection* – this is a complex problem as a whole and requires considerably advanced mathematical competencies. The question may not be very easy for students to understand. The idea that the car producer wants his car to win is rather simple. The complexity is that the new formula has to be valid for all cars, and still make “Ca” the winner. This involves considerable mathematical thinking and argumentation. Students need to identify the relevant mathematical concept of adding weights to different elements within a formula. In this case, students need to understand that the producer wants the strongest features of “Ca” (Safety and Interior) to be weighted most heavily. Plus, it is also desirable if the formula can minimise the stronger points of other cars, especially: External Appearance, and Fuel Efficiency. Using these arguments there are many possible correct answers. An example of a correct answer would be: (5V) + B + O + (5.I).

The question requires students to:

- Reflect on what the numbers in the formula really mean.
- Make the proper choices to weight the different elements within the formula correctly.
- Check the formula for correctness.

Successful students were able to provide a correct rule to make “Ca” the winner. Students were most successful on this question in Japan (45% correct), the partner economy Hong Kong-China (40%), Korea (38%), Belgium (37%) and Switzerland (36%). In seven OECD countries only 20% or fewer students were able to do this successfully.

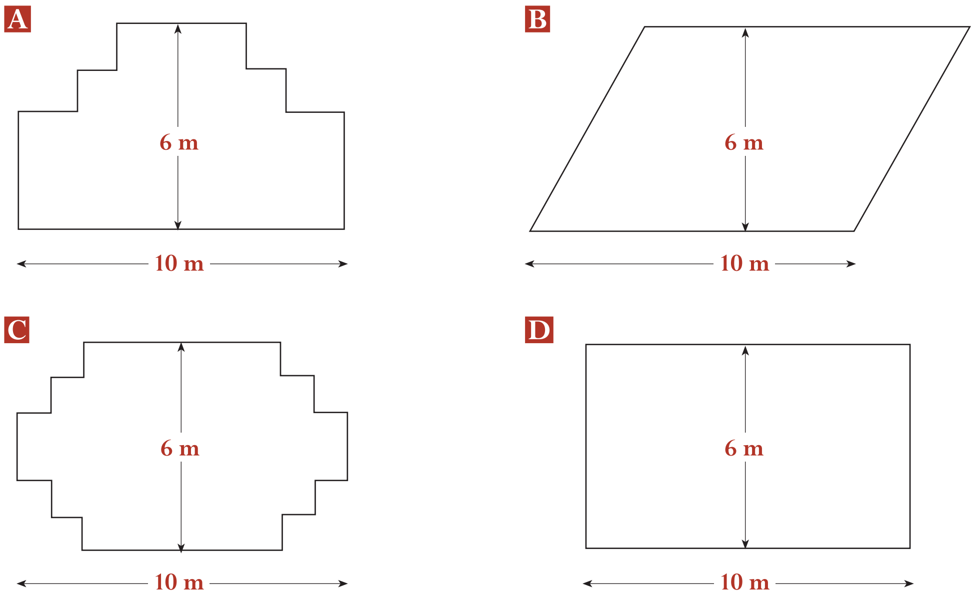
On average across the OECD countries, 19% of students did not respond to this question, but this concerned 32% of students in Denmark and 31% of students in Italy.

Note that this unit includes one other question (THE BEST CAR – Question 1) and this is presented in the section *Examples of easy questions in PISA 2003 mathematics*.

CARPENTER

Question 1: CARPENTER

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



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

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

CARPENTER – Question 1 illustrates Level 6, with a difficulty of 687 PISA score points. On average across OECD countries, 20% of students were able to do this successfully.

Context: *Educational and occupational* – it is the kind of "quasi-realistic" problem that would typically be seen in a mathematics class, rather than being a genuine problem likely to be met in an occupational setting. Whilst not regarded as typical, a small number of such problems have been included in the PISA assessment. However, the competencies needed for this problem are certainly relevant and part of *mathematical literacy*.

Content area: *Space and shape* – geometrical knowledge.

Competency cluster: *Connections* – the problem is certainly non-routine. The students need the competence to recognise that for the purpose of solving the question the two-dimensional shapes A, C and D have the same perimeter; therefore they need to decode the visual information and see similarities and differences. The students need to see whether or not a certain border-shape can be made with 32 metres of timber. In three cases this is rather evident because of the rectangular shapes. But the fourth (B) is a parallelogram, requiring more than 32 metres.

The question requires students to:

- Decode visual information.
- Use argumentation skills.
- Use some technical geometrical knowledge and geometrical insight.
- Use sustained logical thinking.

Successful students answered "Design A, Yes; Design B, No; Design C, Yes; Design D, Yes". Students were most successful on this question in the partner economy Hong Kong-China (40% correct), Japan (38%), Korea (35%) and the partner economy Macao-China (33%). Less than 10% of students were able to do this successfully in Mexico, Greece and the partner countries Tunisia and Brazil. Nearly all students attempted to answer this question with only 2% failing to do so, on average across OECD countries, and this non-response rate did not surpass 5% of students in any of the OECD countries.

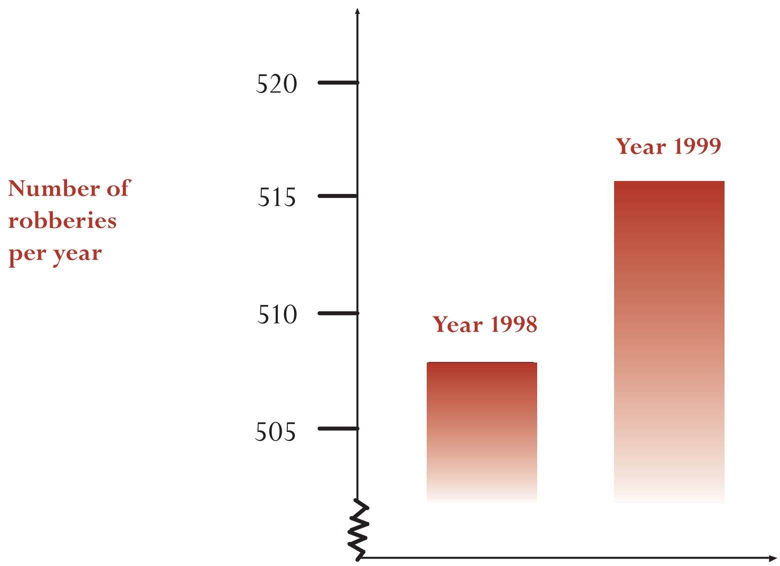
Note: There are actually four questions that students need to answer and this format is often associated with higher question difficulty, since students have to provide the correct response to all parts of the question in order to give a fully correct answer. The sustained logical thinking required to answer all question parts typically indicates a strong understanding of the underlying mathematical issues. On average across OECD countries, 31% of students gave three out of four correct answers. This ranged from 24% of students in Mexico and Turkey to 36% of students in Finland and Denmark. The majority of students across OECD countries tried to answer the question (on average only 2% failed to do so). However, several students had limited success in this. In fact 26% of students on average across OECD countries only gave one out of four correct answers. This was the case for at least 30% of students in Mexico, Greece, Turkey, the United States, Ireland, Portugal and Spain.

ROBBERIES

Question 1: ROBBERIES

A TV reporter showed this graph and said:

"The graph shows that there is a huge increase in the number of robberies from 1998 to 1999."



Do you consider the reporter's statement to be a reasonable interpretation of the graph?

Give an explanation to support your answer.

ROBBERIES – Question 1 illustrates two levels of proficiency in PISA 2003 mathematics depending on whether students give partially or fully correct answers. The latter were discussed in the previous section. Fully correct answers for this question illustrate Level 6, with a difficulty of 694 PISA score points. On average across OECD countries, 30% of students were able to do this successfully.

Context: *Public* – The graph presented in this question was derived from a "real" graph with a similarly misleading message. The graph seems to indicate, as the TV reporter said: "a huge increase in the number of robberies". The students are asked if the statement fits the data It is very important to "look through" data and graphs as they are frequently presented in the media in order to function well in the knowledge society. This constitutes an essential skill in mathematical literacy. (See also the *PISA Assessment Framework 2003*, p. 105). Quite often designers of graphics use their skills (or lack thereof) to let the data support a pre-determined message, often with a political context. This is an example.

Content area: *Uncertainty* – analysis of a graph and interpretation of data. Understanding the issues related to misinterpretation of data. (In this graph the inappropriate cut in the y-axis indicates quite a large increase in the number of robberies, but the absolute difference between the number of robberies in 1998 and 1999 is far from dramatic).

Competency cluster: *Connections* – reasoning and interpretation competencies, together with communication skills.

The question requires students to:

- Understand and decode a graphical representation in a critical way.
- Make judgments and find appropriate argumentation based on mathematical thinking and reasoning (interpretation of data).
- Use some proportional reasoning in a statistical context and a non-familiar situation.
- Communicate effectively their reasoning process.

Successful students indicate that the statement is not reasonable, and explain their judgment in appropriate detail. Their reasoning focuses on the increase of robberies in relative terms and not only on the increase given by an exact number of robberies in absolute terms. Students were most successful on this question giving fully correct answers in Sweden (32% correct), Norway (29%), Finland (27%), Belgium (24%), Italy, New Zealand, Canada and the partner economy Hong Kong-China (23%), Australia and the Netherlands (22%).

Among the OECD countries, the most common type of fully correct answers given by students comprised arguments that the entire graph should be displayed. For example:

- I don't think it is a reasonable interpretation of the graph because if they were to show the whole graph you would see that there is only a slight increase in robberies.
- No, because he has used the top bit of the graph and if you looked at the whole graph from 0 to 520, it wouldn't have risen so much.
- No, because the graph makes it look like there's been a big increase but you look at the numbers and there's not much of an increase.

Such arguments represented at least 70% of the correct answers given in Norway, New Zealand, the United States, Spain, Canada and the United Kingdom.

A significant proportion of fully correct answers given by students also included arguments in terms of the ratio or percentage increase. For example:

- No, not reasonable. ten is not a huge increase compared to a total of 500.
- No, not reasonable. According to the percentage, the increase is only about 2%.
- No. eight more robberies is 1.5% increase. Not much in my opinion!
- No, only eight or nine more for this year. Compared to 507, it is not a large number.

Such arguments represented at least 50% of the correct answers given in Japan, the Czech Republic, Turkey, Italy and Greece, and between 40 and 49% in Austria, France, the Slovak Republic, Switzerland, Portugal, Germany, Poland, Denmark and Ireland.

A minority of the fully correct answers given by students included arguments that trend data are required in order to make such a judgment. For example:

- We cannot tell whether the increase is huge or not. If in 1997, the number of robberies is the same as in 1998, then we could say there is a huge increase in 1999.
- There is no way of knowing what "huge" is because you need at least two changes to think one huge and one small.

While such arguments represented less than 10% of fully correct arguments given in most OECD countries, this was the case for 32% in Korea, 20% in Mexico and 16% in Japan and the Slovak Republic.

Across the OECD countries on average, 15% of students did not respond to this question. This was the case for 30% of students in Greece, 28% in the Slovak Republic and 20% in Turkey, Mexico and Luxembourg, and for between 26 and 35% in the partner countries Serbia, Brazil, Uruguay, the Russian Federation, Tunisia and Indonesia.

CONCLUSION

This chapter illustrated and discussed units and questions of varying difficulties and analysing them in relationship to characteristics that are likely to contribute to making them more or less difficult.

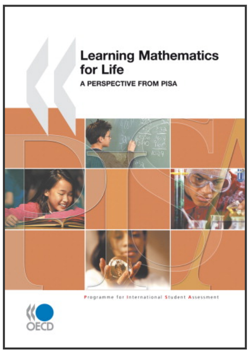
Aspects of the context the question is presented are important in this regard. First, it can be conjectured that contexts that are artificial and that play no role in solving a problem are likely to be less engaging than contexts that both hold more intrinsic interest and are critical to understanding the problem and its solution. On the other hand, contextualised problems that require students to make connections between the problem context and the mathematics needed to solve the problem place a different kind of demand on students. This kind of demand is frequently observed in only the most difficult questions.

Aspects of the question format and particularly of the response requirements are also very important determinants of question difficulty. Questions requiring students to select a response from a number of given options tend to be easier, but this is not always the case, particularly where students must do this a number of times within a single question, *i.e,.* for questions with the complex multiple-choice format. In those questions, a degree of sustained thought is required that exposes the thoroughness of the students' understanding of the mathematical concepts and skills involved in solving the problem.

Questions providing clear direction as to the nature of the answer required, and where convergent thinking is called for to find the one answer that is possible, are usually relatively easy. At the opposite end of the spectrum are questions that require students to construct a response with little or no guidance as to what would constitute an acceptable answer, and where a number of different answers might be acceptable. These questions tend to be more difficult than questions having a more convergent and closed format. When there is an added expectation for students to write an explanation of their conclusion or a justification of their result questions can become very difficult indeed.

Questions with a greater reading load also tend to be more difficult. Sometimes this may be influenced by the extra effort required when more words are involved, but the specific language elements used can also contribute to the level of difficulty. More technical words are less readily handled than simpler words.

Chapters 4 and 5 examine in more detail how students perform on these different types of questions, by analysing their performance on the complete PISA 2003 mathematics question set.



From:

Learning Mathematics for Life

A Perspective from PISA

Access the complete publication at:

https://doi.org/10.1787/9789264075009-en

Please cite this chapter as:

OECD (2010), "A Question of Difficulty: Questions from PISA 2003", in *Learning Mathematics for Life: A Perspective from PISA*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264075009-5-en

OECD