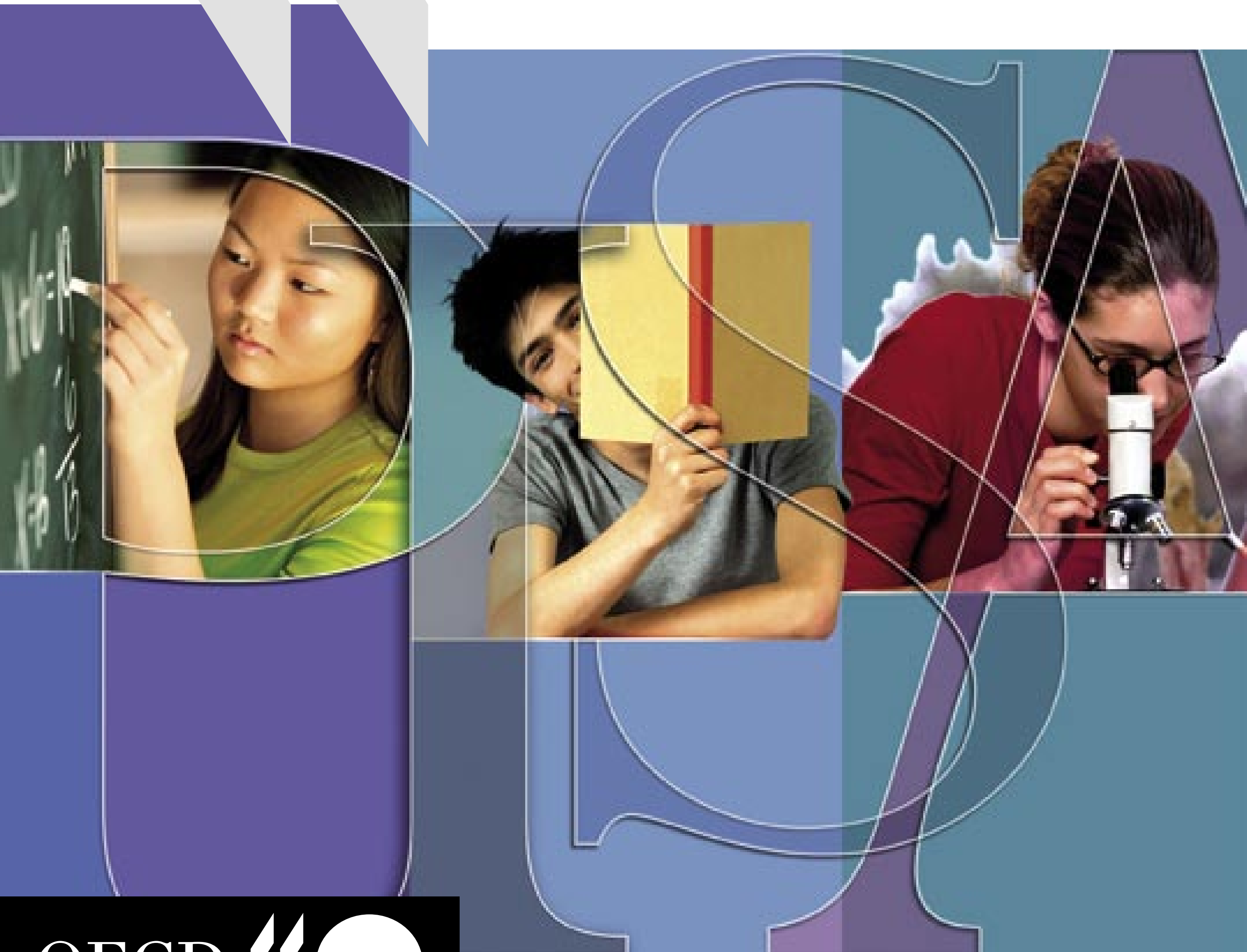
# Assessing Scientific, Reading and Mathematical Literacy

## A FRAMEWORK FOR PISA 2006



OECD

OECD PUBLISHING

Programme for International Student Assessment

**P**rogramme for **I**nternational **S**tudent **A**ssessment

# Assessing Scientific, Reading and Mathematical Literacy

## A Framework for PISA 2006

**OECD**
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

# ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

*This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.*



# Foreword

The OECD Programme for International Student Assessment (PISA), created in 1997, represents a commitment by the governments of OECD member countries to monitor the outcomes of education systems in terms of student achievement, within a common international framework. PISA is, above all, a collaborative effort, bringing together scientific expertise from the participating countries and steered jointly by their governments on the basis of shared, policy-driven interests. Participating countries take responsibility for the project at the policy level. Experts from participating countries also serve on working groups that are charged with linking the PISA policy objectives with the best available substantive and technical expertise in the field of internationally comparative assessment. Through participating in these expert groups, countries ensure that the PISA assessment instruments are internationally valid and take into account the cultural and curricular context of OECD member countries, have strong measurement properties, and place an emphasis on authenticity and educational validity.

PISA 2006 represents a continuation of the data strategy adopted in 1997 by OECD countries. The assessed domains continue to be the same as in 2000 and 2003, however *scientific literacy* is now the major domain and the assessment was carried out using a revised framework. The framework for *reading literacy* remains parallel to the ones used in the 2000 and 2003 assessments and the framework for *mathematical literacy* remains parallel to the ones used in the 2003 assessment and they are respectively presented in the publications *Measuring Student Knowledge and Skills – A New Framework for Assessment* (OECD, 1999) and *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003a).

In a similar way, this new publication presents the guiding principle of the PISA 2006 assessment, which is described in terms of the content that students need to acquire, the processes that need to be performed and the contexts in which knowledge and skills are applied. Further, it illustrates the assessment domains with a range of sample tasks. These have been developed by expert panels under the direction of Raymond Adams, Ross Turner, Barry McCrae and Juliette Mendelovits from the Australian Council for Educational Research (ACER). The science expert group was chaired by Rodger Bybee of the Biological Science Curriculum Study from the United States. The mathematics expert group panel was chaired by Jan de Lange of the University of Utrecht from the Netherlands and the reading expert group was chaired by Irwin Kirsch of Educational Testing Service in the United States until October 2005. After this time John de Jong of the Language Testing Services from the Netherlands became acting chair. The members of the expert groups are listed at the end of this publication. The frameworks have also been reviewed by expert panels in each of the participating countries.

This publication was prepared by the OECD Secretariat, principally by John Cresswell and Sophie Vayssettes. The report is published on the responsibility of the Secretary-General of the OECD.



# Table of Contents









# Introduction

## OVERVIEW

The OECD Programme for International Student Assessment (PISA) is a collaborative effort undertaken by all member countries and a number of non-member partner countries to measure how well students, at age 15, are prepared to meet the challenges they may encounter in future life. Age 15 is chosen because at this age, in most OECD countries, students are approaching the end of compulsory schooling, and so, some measure of the knowledge, skills and attitudes accumulated over approximately ten years of education is gained from an assessment at this time. The PISA assessment takes a broad approach to assessing knowledge, skills and attitudes that reflect current changes in curricula, moving beyond the school based approach towards the use of knowledge in everyday tasks and challenges. The skills acquired reflect the ability of students to continue learning throughout their lives by applying what they learn in school to non-school environments, evaluating their choices and making decisions. The assessment, jointly guided by the participating governments, brings together the policy interests of countries by applying scientific expertise at both national and international levels.

PISA combines the assessment of domain-specific cognitive areas such as science, mathematics and reading with information on students' home background, their approaches to learning, their perceptions of their learning environments and their familiarity with computers. A high priority in PISA 2006 is an innovative assessment of student attitudes towards science – questions about this were contextualised within the cognitive part of the test. Bringing the attitude items closer to the cognitive questions allowed questions to be targeted at specific areas, with the focus on interest in science and students' support for scientific enquiry. Student outcomes are then associated with these background factors.

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

PISA is based on a dynamic model of lifelong learning in which new knowledge and skills necessary for successful adaptation to a changing world are continuously acquired throughout life. PISA focuses on things that 15-year-old students will need in the future and seeks to assess what they can do with what they have learned. The assessment is informed, but not constrained, by the common denominator of national curricula. Thus, while it does assess students' knowledge, PISA also examines their ability to reflect, and to apply their knowledge and experience to real world issues. For example, in order to understand and evaluate scientific advice on food safety an adult would need not only to know some basic facts about the composition of nutrients, but also to be able to apply that information. The term "literacy" is used to encapsulate this broader concept of knowledge and skills.



Box A ■ **What is PISA?**

### Basics

- An internationally standardised assessment that was jointly developed by participating countries and administered to 15-year-olds in educational programmes.
- A survey implemented in 43 countries in the first cycle (32 in 2000 and 11 in 2002), 41 countries in the second cycle (2003) and 56 in the third cycle (2006).
- The test is typically administered to between 4 500 and 10 000 students in each country.

### Content

- PISA 2006 covers the domains of *reading, mathematical* and *scientific literacy* not so much in terms of mastery of the school curriculum, but in terms of important knowledge and skills needed in adult life.
- Emphasis is on the mastery of processes, the understanding of concepts and the ability to function in various situations within each domain.

### Methods

- Paper-and-pencil tests are used, with assessments lasting a total of two hours for each student.
- Test items are a mixture of multiple-choice items and questions requiring students to construct their own responses. The items are organised in groups based on a passage setting out a real-life situation.
- A total of about 390 minutes of test items is covered, with different students taking different combinations of test items.
- Students answer a background questionnaire, which takes 30 minutes to complete, providing information about themselves and their homes. School principals are given a 20-minute questionnaire about their schools.

### Assessment cycle

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

### Outcomes

- A basic profile of knowledge and skills among 15-year-old students
- Contextual indicators relating results to student and school characteristics, with emphasis in 2006 placed on assessing students' attitudes towards science
- Trend indicators showing how results change over time
- A valuable knowledge base for policy analysis and research



PISA is designed to collect information through three-yearly cycles and presents data on the *reading, mathematical* and *scientific literacy* of students, schools and countries. It provides insights into the factors that influence the development of skills and attitudes at home and at school, and examines how these factors interact and what the implications are for policy development.

This publication presents the conceptual framework underlying the PISA 2006 assessments, including a re-developed and expanded framework for *scientific literacy*, incorporating an innovative component on the assessment of students' attitudes towards science, and the frameworks for the assessment of reading and mathematics. Within each domain, the framework defines the contents that students need to acquire, the processes that need to be performed and the contexts in which knowledge and skills are applied. Finally, it illustrates the domain and their aspects with sample tasks.

## BASIC FEATURES OF PISA 2006

PISA 2006 is the third cycle of a data strategy defined in 1997 by participating countries. The publications *Measuring Student Knowledge and Skills – A New Framework for Assessment* (OECD, 1999) and *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003a) presented the conceptual framework underlying the first two cycles of PISA. The results from those cycles were presented in the publications *Knowledge and Skills for Life – First Results from PISA 2000* (OECD, 2001) and *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004), and are also available on the PISA website: *www.pisa.oecd.org*. The results allow national policy makers to compare the performance of their education systems with those of other countries. Similar to the previous cycles, the 2006 assessment covers the domains of *reading, mathematical* and *scientific literacy*, with the major focus on *scientific literacy*. Students also respond to a background questionnaire, and additional supporting information is gathered from the school authorities. Fifty-six countries and regions, including all 30 OECD member countries, are taking part in the PISA 2006 assessment. Together, they comprise almost 90% of the world's economy.

Since the aim of PISA is to assess the cumulative yield of education systems at an age where compulsory schooling is still largely universal, testing focused on 15-year-olds enrolled in both school-based and work-based educational programmes. Between 5 000 and 10 000 students from at least 150 schools will typically be tested in each country, providing a good sampling base from which to break down the results according to a range of student characteristics.

The primary aim of the PISA assessment is to determine the extent to which young people have acquired the wider knowledge and skills in *reading, mathematical* and *scientific literacy* that they will need in adult life. The assessment of cross-curricular competencies continues to be an integral part of PISA 2006. The main reasons for this broadly oriented approach are:

- Although specific knowledge acquisition is important in school learning, the application of that knowledge in adult life depends crucially on the acquisition of broader concepts and skills. In science, having specific knowledge, such as the names of plants and animals, is of less value than understanding broad topics such as energy consumption, biodiversity and human health in thinking about the issues under debate in the adult community. In reading, the capacity to develop interpretations of written material and to reflect on the content and qualities of text are central skills. In mathematics, being able to reason quantitatively and to represent relationships or



dependencies is more apt than the ability to answer familiar textbook questions when it comes to deploying mathematical skills in everyday life.

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

PISA is not a single cross-national assessment of the reading, mathematics and science skills of 15-year-old students. It is an ongoing programme that, over the longer term, will lead to the development of a body of information for monitoring trends in the knowledge and skills of students in various countries as well as in different demographic subgroups of each country. On each occasion, one domain will be tested in detail, taking up nearly two-thirds of the total testing time. The major domain was *reading literacy* in 2000 and *mathematical literacy* in 2003, and is *scientific literacy* in 2006. This will provide a thorough analysis of achievement in each area every nine years and a trend analysis every three.

Similar to previous cycles of PISA, the total time spent on the PISA 2006 tests by each student is two hours, but information is obtained on about 390 minutes worth of test items. The total set of questions is packaged into 13 linked testing booklets. Each booklet is taken by a sufficient number of students for appropriate estimates to be made of the achievement levels on all items by students in each country and in relevant sub-groups within a country (such as males and females, and students from different social and economic contexts). Students also spend 30 minutes answering questions for the context questionnaire.

The PISA assessment provides three main types of outcomes:

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

Although indicators are an adequate means of drawing attention to important issues, they are not usually capable of providing answers to policy questions. PISA has therefore also developed a policy-oriented analysis plan that will go beyond the reporting of indicators.

## WHAT MAKES PISA UNIQUE

PISA is not the first international comparative survey of student achievement. Others have been conducted over the past 40 years, primarily developed by the International Association for the Evaluation of Educational Achievement (IEA) and by the Education Testing Service's International Assessment of Educational Progress (IAEP).



More importantly, these surveys have concentrated on outcomes linked directly to the curriculum and then only to those parts of the curriculum that are essentially common across the participating countries. Aspects of the curriculum unique to one country or a small number of countries have usually not been taken into account in the assessments.

PISA takes a different approach in a number of respects:

- Its *origin:* an initiative taken by governments, whose policy interests the results are addressing.
- Its *regularity:* the commitment to cover multiple assessment domains with updates every three years makes it possible for countries to monitor regularly and predictably their progress in meeting key learning objectives.
- The *age-group covered:* assessing young people near the end of their compulsory schooling gives a useful indication of the performance of education systems. While most young people in OECD countries continue their initial education beyond the age of 15, this is normally close to the end of the initial period of basic schooling in which all young people follow a broadly common curriculum. It is useful to determine, at that stage, the extent to which they have acquired knowledge and skills that will help them in the future, including the individualised paths of further learning they may follow.
- The *knowledge and skills tested:* these are defined not primarily in terms of a common denominator of national school curricula but in terms of what skills are deemed to be essential for future life. This is the most fundamental feature of PISA. School curricula are traditionally constructed largely in terms of bodies of information and techniques to be mastered. They traditionally focus less, within curriculum areas, on the skills to be developed in each domain for use generally in adult life. They focus even less on more general competencies, developed across the curriculum, to solve problems and apply ideas and understanding to situations encountered in life. PISA does not exclude curriculum-based knowledge and understanding, but it tests for it mainly in terms of the acquisition of broad concepts and skills that allow knowledge to be applied. Further, PISA is not constrained by the common denominator of what has been specifically taught in the schools of participating countries.

This emphasis on testing in terms of mastery and broad concepts is particularly significant in light of the concern among nations to develop human capital, which the OECD defines as:

> The knowledge, skills, competencies and other attributes embodied in individuals that are relevant to personal, social and economic well-being.

Estimates of human capital have tended, at best, to be derived using proxies such as level of education completed. When the interest in human capital is extended to include attributes that permit full social and democratic participation in adult life and that equip people to become lifelong learners, the inadequacy of these proxies becomes even clearer.

By directly testing for knowledge and skills close to the end of basic schooling, PISA examines the degree of preparedness of young people for adult life and, to some extent, the effectiveness of education systems. Its ambition is to assess achievement in relation to the underlying objectives (as defined by society) of education systems, not in relation to the teaching and learning of a body of knowledge. This view of educational outcomes is needed if schools and education systems are to be encouraged to focus on modern challenges.



## AN OVERVIEW OF WHAT IS BEING ASSESSED IN EACH DOMAIN

Box B presents a definition of the three domains assessed in PISA 2006. The definitions all emphasise functional knowledge and skills that allow one to participate actively in society. Such participation requires more than just being able to carry out tasks imposed externally by, for example, an employer. It also means being equipped to take part in decision-making processes. In the more complex tasks in PISA, students were asked to reflect on and evaluate material, not just to answer questions that have single correct answers.

### Box B ▪ Definitions of the domains

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

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

***Mathematical literacy:*** An individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen.

*Scientific literacy* (elaborated in Chapter 1) is defined as the ability to use scientific knowledge and processes not only to understand the natural world but to participate in decisions that affect it. *Scientific literacy* is assessed in relation to:

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

*Reading literacy* (elaborated in Chapter 2) is defined in terms of students' ability to understand, use and reflect on written text to achieve their purposes. This aspect of literacy has been well established by previous surveys such as the International Adult Literacy Survey (IALS), but is taken further in PISA by the introduction of an active element – the capacity not just to understand



a text but to reflect on it, drawing on one's own thoughts and experiences. *Reading literacy* is assessed in relation to the:

- *Text format:* Often students' reading assessments have focused on *continuous texts* or prose organised in sentences and paragraphs. PISA introduces in addition *non-continuous texts* that present information in other ways, such as in lists, forms, graphs, or diagrams. It will also distinguish between a range of prose forms, such as narration, exposition and argumentation. These distinctions are based on the principle that individuals will encounter a range of written material in their work-related adult life (*e.g.* application, forms, advertisements) and that it is not sufficient to be able to read a limited number of types of text typically encountered in school.
- *Reading processes (aspects):* Students are not assessed on the most basic reading skills, as it is assumed that most 15-year-old students will have acquired these. Rather, they are expected to demonstrate their proficiency in retrieving information, forming a broad general understanding of the text, interpreting it, reflecting on its contents and reflecting on its form and features.
- *Situations:* These are defined by the use for which the text was constructed. For example, a novel, personal letter or biography is written for people's personal use; official documents or announcements for public use; a manual or report for occupational use; and a textbook or worksheet for educational use. Since some groups may perform better in one reading situation than in another, it is desirable to include a range of types of reading in the assessment items.

*Mathematical literacy* (elaborated in Chapter 3) is concerned with the ability of students to analyse, reason, and communicate ideas effectively as they pose, formulate, solve, and interpret solutions to mathematical problems in a variety of situations. *Mathematical literacy* is assessed in relation to the:

- *Mathematical content:* This is defined mainly in terms of four overarching ideas (*quantity*, *space and shape*, *change and relationships,* and *uncertainty*) and only secondarily in relation to curricular strands such as numbers, algebra and geometry.
- *Mathematical processes:* These are defined by general mathematical competencies. These include the use of mathematical language, modelling and problem-solving skills. Such skills, however, are not separated out in different test items, since it is assumed that a range of competencies will be needed to perform any given mathematical task. Rather, questions are organised in terms of competency clusters defining the type of thinking skill needed.
- *Situations:* These are defined in terms of the ones in which mathematics is used, based on their distance from the students. The framework identifies five situations: personal, educational, occupational, public and scientific.

## ASSESSING AND REPORTING PISA 2006

Similar to the previous assessments in PISA, the assessment in 2006 consists of pencil and paper instruments for reasons of feasibility. The assessment includes a variety of types of questions. Some require students to select or produce simple responses that can be directly compared with a single correct answer, such as multiple-choice or closed-constructed response items. These questions have either a correct or incorrect answer and often assess lower-order skills. Others are more constructive, requiring students to develop their own responses designed to measure broader constructs than those captured by more traditional surveys, allowing for a wider range of acceptable responses and more complex marking that can include partially correct responses.



Not all students answer all questions in the assessment. The PISA 2006 test units are arranged in 13 clusters, with each cluster designed to occupy 30 minutes of testing time. There are seven science clusters, two reading clusters and four mathematics clusters. The clusters are placed in 13 booklets, according to a rotated test design. Each booklet contains four clusters and each student is assigned one of these two-hour booklets. There is at least one science cluster in each booklet.

Literacy in PISA is assessed through units consisting of a stimulus (*e.g.* text, table, chart, figures etc.) followed by a number of tasks associated with this common stimulus. This is an important feature, allowing questions to go into greater depth than they could if each question introduced a wholly new context. It allows time for the student to digest material that can then be used to assess multiple aspects of performance.

Results from PISA have been reported using scales with an average score of 500 and a standard deviation of 100 for all three domains, which means that two-thirds of students across OECD countries scored between 400 and 600 points. These scores represent degrees of proficiency in a particular aspect of literacy. As *reading literacy* was the major domain in 2000, the reading scales were divided into five levels of knowledge and skills. The main advantage of this approach is that it describes what students can do by associating the tasks with levels of difficulty. Additionally, results were also presented through three subscales of reading: retrieving information, interpreting texts, and reflection and evaluation. A proficiency scale was also available for *mathematical* and *scientific literacy*, though without levels thus recognising the limitation of the data from minor domains. PISA 2003 built upon this approach by specifying six proficiency levels for the mathematical literacy scale, following a similar approach to what was done in reading. There were four subscales in *mathematical literacy: space and shape, change and relationships, quantity* and *uncertainty*. The reporting of *scientific literacy* will be in a similar manner and will also present results in different areas. PISA 2003 offered the first opportunity to present trend results for *reading, mathematical* and *scientific literacy* and the results from PISA 2006 will provide extra information for this analysis.

## THE CONTEXT QUESTIONNAIRES AND THEIR USE

To gather contextual information, PISA asks students and the principals of their schools to respond to background questionnaires of around 30 minutes in length. These questionnaires are central to the analysis of results in terms of a range of student and school characteristics. The questionnaires from PISA 2000 and 2003 are available on the PISA website: *www.pisa.oecd.org.*

The questionnaires seek information about:

- Students and their family backgrounds, including their economic, social and cultural capital
- Aspects of students' lives, such as their attitudes towards learning, their habits and life inside school, and their family environment
- Aspects of schools, such as the quality of the schools' human and material resources, public and private control and funding, decision-making processes, and staffing practices
- Context of instruction, including institutional structures and types, class size, and the level of parental involvement



- Strategies of self-regulated learning, motivational preferences and goal orientations, self-related cognition mechanisms, action control strategies, preferences for different types of learning situations, learning styles, and social skills required for co-operative or competitive learning
- Aspects of learning and instruction in science, including students' motivation, engagement and confidence with science, and the impact of learning strategies on achievement related to the teaching and learning of science

Two additional questionnaires are offered as international options:

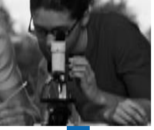
- A computer familiarity questionnaire focusing on: *i)* availability and use of information and communications technology (ICT), including the location where ICT is mostly used as well as the type of use; *ii)* ICT confidence and attitudes, including self-efficacy and attitudes towards computers; and *iii)* learning background of ICT, focusing on where students learned to use computers and the Internet. The OECD published a report resulting from analysis of data collected via this questionnaire in 2003, *Are Students Ready for a Technology-Rich World? What PISA Studies Tell Us* (OECD, 2005).
- A parent questionnaire focusing on a number of topics including the student's past science activities, parents' views on the student's school, parents' views on science in the student's intended career and the need for scientific knowledge and skills in the job market, parents' views on science and the environment, the cost of education services, and parents' education and occupation.

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

## COLLABORATIVE DEVELOPMENT OF PISA AND ITS ASSESSMENT FRAMEWORKS

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

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





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

The design and implementation of the present surveys, within the framework established by the PGB, is the responsibility of an international consortium led by the Australian Council for Educational Research (ACER). Other partners in this consortium include the National Institute for Educational Measurement (CITO) in the Netherlands, WESTAT and the Educational Testing Service (ETS) in the United States, and the National Institute for Educational Policy Research (NIER) in Japan.

The OECD Secretariat has overall managerial responsibility for the programme, monitors its implementation on a day-to-day basis, acts as the secretariat for the PGB, builds consensus among countries and serves as the interlocutor between the PGB and the international consortium charged with implementation. The OECD Secretariat is also responsible for the production of the indicators, and the analysis and preparation of the international reports and publications in co-operation with the PISA consortium, in close consultation with member countries both at the policy level (PGB) and at the implementation level (National Project Managers).

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

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

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

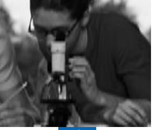
- A framework provides a common language and a vehicle for discussing the purpose of the assessment and what it is trying to measure. Such a discussion encourages the development of a consensus around the framework and the measurement goals.
- An analysis of the kinds of knowledge and skills associated with successful performance provides a basis for establishing standards or levels of proficiency. As the understanding of what is being measured and the ability to interpret scores along a particular scale evolve, an empirical basis for communicating a richer body of information to various constituencies can be developed.
- Identifying and understanding particular variables that underlie successful performance further the ability to evaluate what is being measured and to make changes to the assessment over time.



The understanding of what is being measured and its connection to what we say about students provides an important link between public policy, assessment and research which, in turn, enhances the usefulness of the data collected.

1

# Scientific Literacy





## INTRODUCTION

The assessment of *scientific literacy* has particular importance in PISA 2006, where it is the major domain being assessed. Since this is the first time that it is being tested in such detail, the domain has undergone considerable development since the 2003 survey, with an expanded interpretation of what is being assessed. This involves not only a more detailed description of *scientific literacy,* but also an important innovation in the approach to assessment that has relevance for all of PISA in the future. For the first time, the main assessment instrument includes questions on attitudes alongside the testing of cognitive abilities and knowledge. By exploring the extent to which the issues that they are addressing in the course of the test excite students' interest, this strengthens the assessment of the attitudinal and motivational characteristics that will be important to their future engagement with science. Previously, questions about these aspects have been limited to a separate questionnaire asking in more general terms about aspects such as interest and motivation.

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

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

As the basis for an international assessment of 15-year-old students, it seems reasonable, therefore, to ask: "What is it important for citizens to know, value, and be able to do in situations involving science and technology?" Answering this question establishes the basis for an assessment of students in these respects: their knowledge, values and abilities today relate to what is needed in the future. Central to the answer are the competencies that lie at the heart of the PISA 2006 definition of *scientific literacy*. These ask how well students:

- Identify scientific issues
- Explain phenomena scientifically
- Use scientific evidence

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

The issue of identifying what citizens should know, value and be able to do in situations involving science and technology, seems simple and direct. However doing so raises questions about scientific



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

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

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

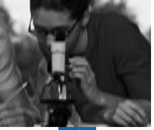
Can citizens distinguish claims that are scientifically sound from those that are not? Ordinary citizens are generally not called on to judge the worth of major theories or potential advances in science. But they do make decisions based on the facts in advertisements, evidence in legal matters, information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

## DEFINITION OF THE DOMAIN

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

Such goals provide an orientation and emphasis for the science education of all people (Fensham, 1985). The competencies assessed in PISA 2006 are broad and include aspects that relate to personal utility, social responsibility, and the intrinsic and extrinsic value of scientific knowledge.

The above discussion frames a central point of the PISA 2006 science assessment: The assessment should focus on competencies that clarify what 15-year-old students know, value and are able to do within reasonable and appropriate personal, social and global contexts. This perspective differs from one grounded exclusively in school science programmes and extensively based only on the disciplines





of science; but it includes problems situated in educational contexts and also in professional ones, and recognises the essential place of the knowledge, methods, attitudes, and values that define scientific disciplines. The term that best describes the overall purposes of the PISA 2006 science assessment is *scientific literacy* (Bybee, 1997b; Fensham, 2000; Graber and Bolte, 1997; Mayer, 2002; Roberts, 1983; UNESCO, 1993).

PISA 2006 aims to assess both the cognitive and affective aspects of students' *scientific literacy*. The cognitive aspects include students' knowledge and their capacity to use this knowledge effectively, as they carry out certain cognitive processes that are characteristic of science and scientific enquiries of personal, social, or global relevance. In assessing scientific competencies, PISA is concerned with issues to which scientific knowledge can contribute and which will involve students, either now or in the future, in making decisions. From the point of view of their scientific competencies, students respond to such issues in terms of their understanding of relevant scientific knowledge, their ability to access and evaluate information, their ability to interpret evidence bearing on the issue and their ability to identify the scientific and technological aspects of the issue (Koballa, Kemp and Evans, 1997; Law, 2002). PISA also assesses non-cognitive aspects: how students respond affectively. Attitudinal aspects of their response engage their interest, sustain their support, and motivate them to take action (Schibeci, 1984). Through such considerations we are led to define the overarching domain of *scientific literacy* for PISA 2006.

### Box 1.1 ■ **Scientific knowledge: PISA 2006 terminology**

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

The term *scientific literacy* has been chosen because it is recognised as representing the goals of science education that should apply to all students, connotes a broadness and an applied nature to the purposes of science education, represents a continuum of scientific knowledge and the cognitive abilities associated with scientific enquiry, incorporates multiple dimensions, and includes the relationships between science and technology. Together, the scientific competencies at the heart of the definition characterise a foundation for *scientific literacy,* and the objective of the PISA 2006 science assessment – to assess the degree to which the competencies have been developed (Bybee, 1997a; Fensham, 2000; Law, 2002; Mayer and Kumano, 2002).



**Box 1.2 ■ PISA 2006 *Scientific literacy***

For the purposes of PISA 2006, *scientific literacy*[2] refers to an individual's:

- Scientific knowledge and use of that knowledge to identify questions, acquire new knowledge, explain scientific phenomena and draw evidence-based conclusions about science-related issues
- Understanding of the characteristic features of science as a form of human knowledge and enquiry
- Awareness of how science and technology shape our material, intellectual, and cultural environments
- Willingness to engage in science-related issues and with the ideas of science, as a reflective citizen

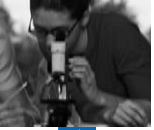
The following remarks further clarify this definition.

### Scientific literacy

Using the term "scientific literacy" rather than "science" underscores the importance that the PISA 2006 science assessment places on the application of scientific knowledge in the context of life situations, compared with the simple reproduction of traditional school science knowledge. The functional use of knowledge requires the application of those processes that are characteristic of science and scientific enquiry (the scientific competencies) and is regulated by the individual's appreciation, interest, values, and action relative to scientific matters. A student's ability to carry out the scientific competencies involves both knowledge of science and an understanding of the characteristics of science as a way of acquiring knowledge (*i.e.* knowledge about science). The definition also recognises that the disposition to carry out these competencies depends upon an individual's attitudes towards science and a willingness to engage in science-related issues. Note that non-cognitive aspects such as motivation are themselves considered to be competencies.

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

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





to draw definite conclusions, thus making it necessary to speculate, cautiously and consciously, about the information that is available.

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

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

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

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

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

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

Inevitably, *scientific literacy* draws upon *reading* and *mathematical literacies* (Norris and Phillips, 2003). For example, *reading literacy* is necessary when a student is demonstrating an understanding of scientific terminology. Similarly, aspects of *mathematical literacy* are required in data interpretation contexts. The intersection of these other literacies with the PISA 2006 definition and assessment of scientific literacy cannot be avoided; however, at the core of each assessment task there should be aspects that are unambiguously *scientific literacy*.

Compared to the definition of *scientific literacy* for PISA in 2000 and 2003, the definition for 2006 has been elaborated and enhanced. For the previous two assessments, when science was a minor domain, *scientific literacy* was defined as follows:



*Scientific literacy* is the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity. (OECD, 1999, 2000, 2003a)

The initial assertions of the 2000, 2003 and 2006 definitions are fundamentally the same in that they centre on individuals' uses of scientific knowledge to draw conclusions. While the 2000 and 2003 definition embedded knowledge of science and understandings about science within the terms of scientific knowledge, the 2006 definition separates and elaborates this aspect of *scientific literacy* through the addition of terms that underscore students' knowledge about the characteristic features of science. Both definitions then refer to the application of scientific knowledge to understand, and ultimately to make informed decisions about, the natural world. In PISA 2006, this part of the definition is enhanced by the addition of knowledge of the relationship between science and technology – an aspect of *scientific literacy* that was assumed but not elaborated in the earlier definition. In today's world, science and technology are closely linked, often having synergistic relationships with each other.

In contrast to the earlier definition, the PISA 2006 definition of *scientific literacy* has been expanded by explicitly including attitudinal aspects of students' responses to issues of scientific and technological relevance. In summary, the 2006 definition is conceptually in accord with the 2000 and 2003 definition, with the exception of the addition of attitudinal responses. However the attitudinal element is reported separately and therefore does not impact on the comparability of the cognitive aspect over time. Other changes, for example elaborating knowledge about science, and science-based technology, represent an increased emphasis on particular aspects that were embedded or assumed in the earlier definition.

## ORGANISATION OF THE DOMAIN

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

For purposes of assessment, the PISA 2006 definition of *scientific literacy* may be characterised as consisting of four interrelated aspects:

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

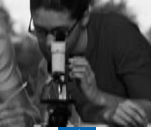
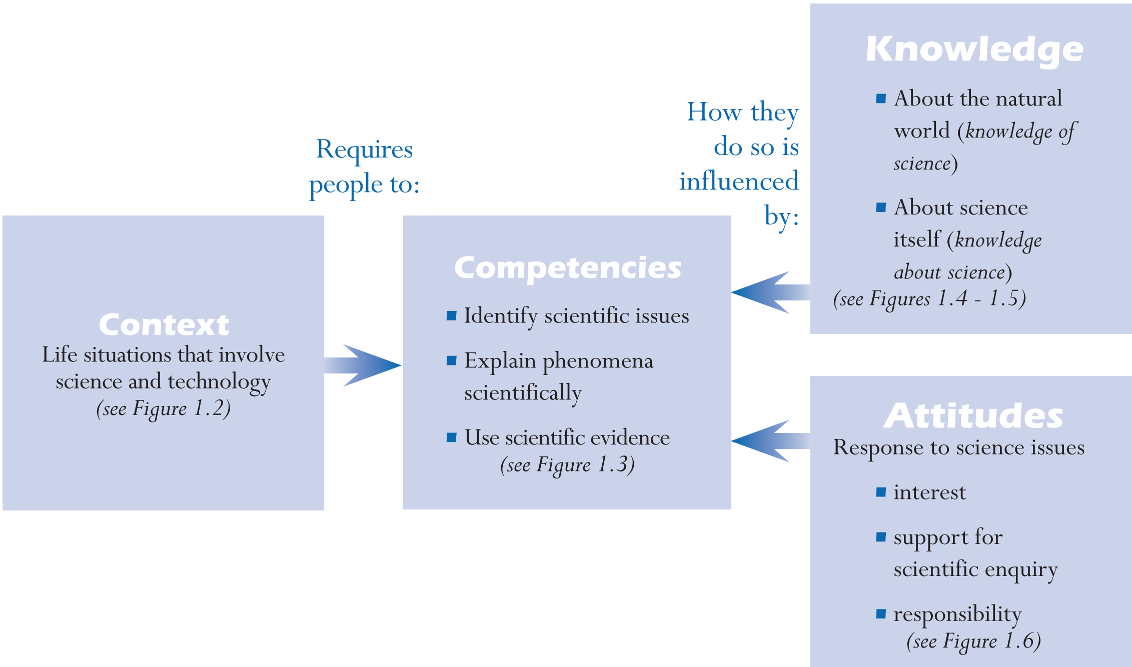




Figure 1.1 ■ **Framework for PISA 2006 science assessment**




The following sections restate and elaborate the interrelated aspects of *scientific literacy*. In highlighting these aspects, the PISA 2006 scientific literacy framework has ensured that the focus of the assessment is upon the outcomes of science education as a whole. Several questions, based on the PISA 2006 perspective of *scientific literacy* lie behind the organisation of this section of the framework. They are:

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

## SITUATIONS AND CONTEXT

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

The situation is the part of the student's world in which the tasks are placed. Assessment items are framed in situations of general life and not limited to life in school. In the PISA 2006 science assessment, the focus of the items is on situations relating to the self, family and peer groups (*personal*), to the community (*social*) and to life across the world (*global*). A further type of situation, appropriate to some topics, is the *historical* one, in which understanding of the advances in scientific knowledge can be assessed.



The context of an item is its specific setting within a situation. It includes all the detailed elements used to formulate the question.

PISA 2006 assesses important scientific knowledge relevant to the science education curricula of participating countries without being constrained to the common aspects of participants' national curricula. The assessment does this by requiring evidence of the successful use of scientific competencies in important situations reflecting the world and in accordance with PISA's focus on *scientific literacy*. This, in turn, involves the application of selected knowledge about the natural world, and about science itself, and evaluation of students' attitudes towards scientific matters.

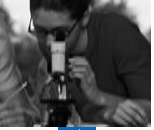
Figure 1.2 lists the applications of science, within *personal, social,* and *global* situations, primarily used as the contexts for assessment exercises. However, other situations (*e.g. technological, historical*) and areas of application are used. The applications were drawn from a wide variety of life situations and were generally consistent with the areas of application for *scientific literacy* in the 2000 and 2003 PISA frameworks. The areas of application are: "health", "natural resources", "the environment", "hazards", and "the frontiers of science and technology". They are the areas in which *scientific literacy* has particular value for individuals and communities in enhancing and sustaining quality of life, and in the development of public policy.

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

The contexts used for assessment items are chosen in the light of relevance to students' interests and lives. Science items are developed keeping in mind linguistic and cultural differences in participating countries.

Figure 1.2 ■ **Contexts for the PISA 2006 science assessment**

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





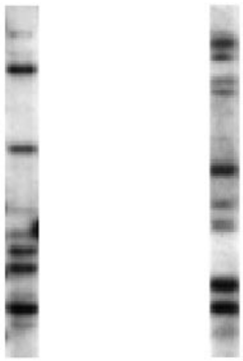
Science Example 1 is part of a unit titled *CATCHING THE KILLER*. The stimulus material is a newspaper article that establishes the context for the unit. The area of application is "Frontiers of science and technology" within a social setting.

## Science Example 1: CATCHING THE KILLER

**DNA TO FIND KILLER**

**Smithville, yesterday:** A man died from multiple stab wounds in Smithville today. Police say that there were signs of a struggle and that some of the blood found at the scene of the crime did not match the victim's blood. They believe that this blood came from the killer.

To help find the killer, police scientists have prepared a DNA profile from the blood sample. When compared to DNA profiles of convicted criminals, kept on a computer database, no match was found.



**Person A** **Person B**

Photo of typical DNA profiles from two people. The bars are different fragments of each person's DNA. Each person has a different pattern of bars. Like fingerprints, these patterns can identify a person.

**Police have now arrested a local man seen arguing with the victim earlier in the day. They have applied for permission to collect a sample of the suspect's DNA.**

Sergeant Brown of the Smithville police said, "We just need to take a harmless scraping from the inside of the cheek. From this scraping scientists can extract DNA and form a DNA profile like the ones pictured."

Except for identical twins, there is only a 1 in 100 million chance that two people will have the same DNA profile.

### Question 1: CATCHING THE KILLER

This newspaper article refers to the substance DNA. What is DNA?

A. A substance in cell membranes that stops the cell contents leaking out.

B. A molecule that contains the instructions to build our bodies.

C. A protein found in blood that helps carry oxygen to our tissues.

D. A hormone in blood that helps regulate glucose levels in body cells.

### Question 2: CATCHING THE KILLER

Which one of the following questions *cannot* be answered by scientific evidence?

A. What was the medical or physiological cause of the victim's death?

B. Who was the victim thinking of when he died?

C. Is taking cheek scrapings a safe way to collect DNA samples?

D. Do identical twins have exactly the same DNA profile?



## SCIENTIFIC COMPETENCIES

The PISA 2006 science assessment gives priority to the competencies listed in Figure 1.3, the ability to: identify scientifically-oriented issues; describe, explain or predict phenomena based on scientific knowledge; interpret evidence and conclusions; and use scientific evidence to make and communicate decisions. These competencies involve scientific knowledge – both knowledge of science and knowledge about science itself as a form of knowledge and an approach to enquiry.

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

Justification for an emphasis on the scientific competencies of Figure 1.3 in PISA 2006 rests on the importance of these competencies for scientific investigation. They are grounded in logic, reasoning, and critical analysis. An elaboration of the scientific competencies follows.

Figure 1.3 ■ **PISA 2006 scientific competencies**

**Identifying scientific issues**

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

**Explaining phenomena scientifically**

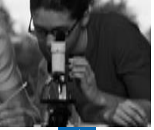
- Applying knowledge of science in a given situation
- Describing or interpreting phenomena scientifically and predicting changes
- Identifying appropriate descriptions, explanations, and predictions

**Using scientific evidence**

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

### Identifying scientific issues

It is important to be able to distinguish scientific issues and content from other forms of issues. Importantly, scientific issues must lend themselves to answers based on scientific evidence. The competency *identifying scientific issues* includes recognising questions that it would be possible to investigate scientifically in a given situation and identifying keywords to search for scientific





information on a given topic. It also includes recognising key features of a scientific investigation: for example, what things should be compared, what variables should be changed or controlled, what additional information is needed, or what action should be taken so that relevant data can be collected.

*Identifying scientific issues* requires students to possess knowledge about science itself, but may also draw, to varying degrees, on their knowledge of science. Question 2 of *CATCHING THE KILLER* (Science Example 1) requires students to identify a question that cannot be investigated scientifically. The item mainly assesses students' knowledge of what types of questions can be investigated scientifically (Knowledge about science, category: "Scientific enquiry"), but assumes a knowledge of science (category: "Living systems") that 15-year-old students could be expected to possess.

### Explaining phenomena scientifically

Students demonstrate the competency *explaining phenomena scientifically* by applying appropriate knowledge of science in a given situation. The competency includes describing or interpreting phenomena and predicting changes, and may involve recognising or identifying appropriate descriptions, explanations, and predictions. Question 1 of *CATCHING THE KILLER* (Science Example 1) requires students to draw on their knowledge of science (category: "Living systems") to recognise the appropriate description of DNA.

### Using scientific evidence

The competency *using scientific evidence* requires students to make sense of scientific findings as evidence for claims or conclusions. The required response can involve knowledge about science or knowledge of science or both. The question in *MALARIA* (Science Example 2) requires students to make conclusions based on the evidence presented about the life cycle of a mosquito. The item mainly assesses whether students can interpret a standard representation (model) of a life cycle – this is knowledge about science (category: "Scientific explanations" – see Figure 1.5).

*Using scientific evidence* includes accessing scientific information and producing arguments and conclusions based on scientific evidence (Kuhn, 1992; Osborne, Erduran, Simon and Monk, 2001). The competency may also involve: selecting from alternative conclusions in relation to evidence; giving reasons for or against a given conclusion in terms of the process by which the conclusion was derived from the data provided; and identifying the assumptions made in reaching a conclusion. Reflecting on the societal implications of scientific or technological developments is another aspect of this competency.

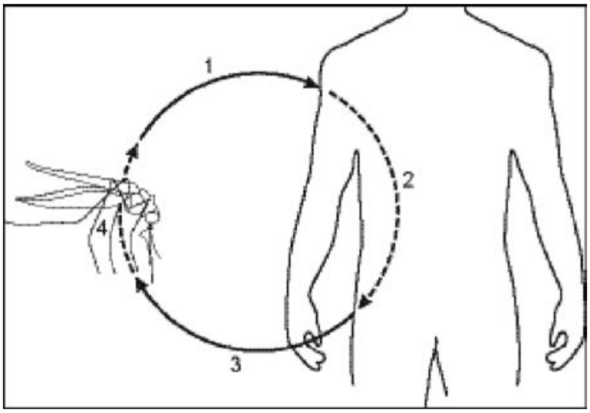
Students may be required to express their evidence and decisions, through their own words, diagrams or other representations as appropriate, to a specified audience. In short, students should be able to present clear and logical connections between evidence and conclusions or decisions.

## Science Example 2: MALARIA

Malaria is the cause of more than one million deaths every year. The fight against malaria is currently in crisis. Mosquitoes pass the malaria parasite from person to person. The malaria-carrying mosquito has become resistant to many pesticides. Also, medicines against the malaria parasite are getting less and less effective.

Life cycle of the malaria parasite

**1**
The malaria parasite is passed to a human through a female mosquito bite or sting.




**2**
The parasite reproduces in the human body and causes the first symptoms of the disease.

**3**
The parasite is transferred to another mosquito that bites or stings the infected person.

**4**
The parasite also reproduces in the mosquito's body, but the mosquito is not affected.

### Question 1: MALARIA

Three methods of preventing the spread of malaria are given below.

Which of the stages (1, 2, 3 and 4) in the life cycle of a malaria parasite are *directly* affected by each method? *Circle the relevant stage(s) for each method (more than one stage may be affected by a single method).*

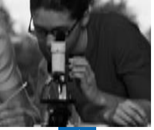
| Method of preventing the spread of malaria | Stages in the life cycle of the parasite that are affected |
|---|---|
| Sleeping underneath a mosquito net. | 1 2 3 4 |
| Taking medicines against malaria. | 1 2 3 4 |
| Using pesticides against mosquitoes. | 1 2 3 4 |

## SCIENTIFIC KNOWLEDGE

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

### Knowledge of science

Given that only a sample of students' knowledge of science can be assessed in the PISA 2006 science assessment, it is important that clear criteria are used to guide the selection of knowledge that is assessed. Moreover, the objective of PISA is to describe the extent to which students can apply their knowledge in contexts of relevance to their lives. Accordingly, the assessed knowledge





will be selected from the major fields of physics, chemistry, biology, Earth and space science, and technology[3] according to the following criteria:

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

Figure 1.4 shows the *knowledge of science* categories and examples of content selected by applying these criteria. This knowledge is required for understanding the natural world and for making sense

Figure 1.4 ■ **PISA 2006 categories of *knowledge of science***

**Physical systems**

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

**Living systems**

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

**Earth and space systems**

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

**Technology systems**

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



of experiences in *personal, social* and *global* contexts. For these reasons the framework uses the term "systems" instead of "sciences" in the descriptors of the major fields. The intention is to convey the idea that citizens have to understand concepts from the physical and life sciences, Earth and space science, and technology, in several different contexts.

The examples listed in Figure 1.4 convey the meanings of the categories; there is no attempt to list comprehensively all the knowledge that could be related to each of the knowledge of science categories. Question 1 of *CATCHING THE KILLER* (Science Example 1) assesses students' *knowledge of science* in the category "Living systems".

### Knowledge about science

Figure 1.5 displays the categories and examples of content for *knowledge about science.* The first category, "Scientific enquiry," centres on enquiry as the central process of science and the various components of that process. The second category, closely related to enquiry, is "Scientific explanations". Scientific explanations are the results of scientific enquiry. One can think of enquiry as the means of science (how scientists get data) and explanations as the goals of science (how scientists use data). The examples listed in Figure 1.5 convey the general meanings of the categories; there is no attempt to list comprehensively all the knowledge that could be related to each category.

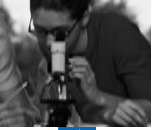
Figure 1.5 ■ **PISA 2006 categories of *knowledge about science***

**Scientific enquiry**

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

**Scientific explanations**

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





Science Example 3 is part of a unit titled *SCHOOL MILK STUDY*, with a historical setting and health as the area of application. Both questions assess students' knowledge about science, in the category "Scientific enquiry". Question 1 requires students to identify the possible purposes of the study (competency: "Identifying scientific issues"). The competency classification of Question 2 is also "Identifying scientific issues" (rather than "Using scientific evidence") since the most obvious assumption (that the three groups of students were not significantly different in any relevant way) relates to the design of the study.

## Science Example 3: SCHOOL MILK STUDY

In 1930, a large-scale study was carried out in the schools in a region of Scotland. For four months, some students received free milk and some did not. The head teachers in each school chose which of their students received milk. Here is what happened:

- 5 000 school children received an amount of unpasteurised milk each school day
- Another 5 000 school children received the same amount of pasteurised milk
- 10 000 school children did not receive any milk at all

All 20 000 children were weighed and had their heights measured at the beginning and the end of the study.

### Question 1: SCHOOL MILK STUDY

Is it likely that the following questions were research questions for the study?

*Circle "Yes" or "No" for each question.*

| Is it likely that this was a research question for the study? | Yes or No? |
|---|---|
| What has to be done to pasteurise milk? | Yes / No |
| What effect does the drinking of additional milk have on school children? | Yes / No |
| What effect does milk pasteurisation have on school children's growth? | Yes / No |
| What effect does living in different regions of Scotland have on school children's health? | Yes / No |

### Question 2: SCHOOL MILK STUDY

On average, the children who received milk during the study gained more in height and weight than the children who did not receive milk.

One possible conclusion from the study, therefore, is that school children who drink a lot of milk grow faster than those who do not drink a lot of milk.

To have confidence in this conclusion, indicate one assumption that needs to be made about these two groups of students in the study.





## ATTITUDES TOWARDS SCIENCE

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

The PISA 2006 assessment of science takes an innovative approach to assessing student attitudes. Not only does it ask them about what they think about science in the student questionnaire, but it also asks them, in the course of the science part of the assessment, what their attitudes are towards the issues that they are being tested on.

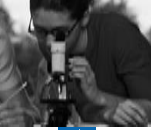
The survey's attention to attitudes towards science is based on the belief that a person's *scientific literacy* includes certain attitudes, beliefs, motivational orientations, sense of self-efficacy, values, and ultimate actions. The inclusion of attitudes and the specific areas selected for PISA 2006 is supported by and builds upon Klopfer's (1976) structure for the affective domain in science education as well as reviews of attitudinal research (for example, Gardner, 1975, 1984; Gauld and Hukins, 1980; Blosser, 1984; Laforgia, 1988; Schibeci, 1984).

The PISA 2006 science assessment evaluated students' attitudes in three areas: *interest in science*, *support for scientific enquiry* and *responsibility towards resources and environments* (see Figure 1.6). These areas were selected because they provide an international portrait of students' general appreciation of science, their specific scientific attitudes and values, and their responsibility towards selected science-related issues that have national and international ramifications. This was not an assessment of students' attitudes toward school science programs or teachers. The results may provide information about the emerging problem of declining enrolments for science studies among young people.

*Interest in science* was selected because of its established relationships with achievement, course selection, career choice, and lifelong learning. The relationship between (individual) interest in science and achievement has been the subject of research for more than 40 years although there is still debate about the causal link (see, for example, Baumert and Köller, 1998; Osborne, Simon & Collins, 2003). The PISA 2006 science assessment addressed students' interest in science through knowledge about their engagement in science-related social issues, their willingness to acquire scientific knowledge and skills, and their consideration of science-related careers.

*Support for scientific enquiry* is widely regarded as a fundamental objective of science education and as such warrants assessing. It is a similar construct to "adoption of scientific attitudes" as identified by Klopfer (1971). Appreciation of and support for scientific enquiry implies that students value scientific ways of gathering evidence, thinking creatively, reasoning rationally, responding critically, and communicating conclusions as they confront life situations related to science. Aspects of this area in PISA 2006 include the use of evidence (knowledge) in making decisions, and the appreciation for logic and rationality in formulating conclusions.

*Responsibility towards resources and environments* is of international concern, as well as being of economic relevance. Attitudes in this area have been the subject of extensive research since the 1970s (see, for example, Bogner and Wiseman, 1999; Eagles & Demare, 1999; Weaver, 2002; Rickinson, 2001). In December 2002, the United Nations approved resolution 57/254 declaring





the ten-year period beginning on 1 January 2005 to be the United Nations Decade of Education for Sustainable Development (UNESCO, 2003). The International Implementation Scheme (UNESCO, September 2005) identifies environment as one of the three spheres of sustainability (along with society (including culture) and economy) that should be included in all education for sustainable development programmes.

PISA 2006 gathers data about such student attitudes both by posing questions in the student questionnaire and in contextualised test items – that is questions about attitudes towards issues posed immediately after test questions related to these issues (see Box 1.2). The student questionnaire collects information about students' attitudes in all three areas: *interest in science*, *support for scientific enquiry* and *responsibility towards resources and environments,* in a non-contextualised manner. Additional data concerning students' engagement in science (*e.g.* self-efficacy, enjoyment of science and frequency of out of school scientific activities) were also collected via the student questionnaire, as were students' views on the value of science for their own lives (*e.g.* further education and career choices) and for society (*e.g.* social and economic benefits).

Contextualised items are used in relation to interest in learning about science and student support for scientific enquiry. Contextualised items add value to the assessment in that they provide data on whether students' attitudes differ when assessed in or out of context, whether they vary between contexts, and whether they correlate with performance at the unit level. One aspect of students' *interest in science* (namely, their interest in learning about science) and of students' *support for scientific enquiry* was assessed in the test using embedded items that targeted *personal, social,* and *global* issues.

The results of PISA 2006 will provide important information for educational policy makers in the participating countries. The combined richness of the data obtained through both the student questionnaire and the embedded attitudinal items should generate new knowledge about students' predispositions towards scientifically literate behaviours. Further, since the literature contains conflicting reports on the correlation between attitudes and performance in science, it remains to be seen how student attitudinal data (concerning students' *interest in science*, *support for scientific enquiry* and *responsibility towards resources and environments*), collected via the test and the questionnaire, correlates with student performance. Other data obtained from the student questionnaire, such as students' engagement in science and science-related behaviours, also will be reported and linked with student performance.

## ASSESSING SCIENTIFIC LITERACY

### Test characteristics

In accordance with the PISA definition of *scientific literacy,* test questions (items) require the use of the scientific competencies (see Figure 1.3) within a context (see Figure 1.2). This involves the application of scientific knowledge (see Figures 1.4 and 1.5) and reflects aspects of the respondents' attitudes towards scientific matters (see Figure 1.6).

Figure 1.7 is a variation of Figure 1.1 that presents the basic components of the PISA framework for the 2006 scientific literacy assessment in a way that can be used to relate the framework with the structure and the content of assessment units. Figure 1.7 may be used both synthetically as a tool to plan assessment exercises, and analytically as a tool to study the results of standard assessment exercises. As a starting point to construct assessment units, we could consider the contexts that

Figure 1.6 ■ **PISA 2006 areas for assessment of attitudes**

**Interest in science**

- Indicate curiosity in science and science-related issues and endeavours
- Demonstrate willingness to acquire additional scientific knowledge and skills, using a variety of resources and methods
- Demonstrate willingness to seek information and have an ongoing interest in science, including consideration of science-related careers

**Support for scientific enquiry**

- Acknowledge the importance of considering different scientific perspectives and arguments
- Support the use of factual information and rational explanations
- Express the need for logical and careful processes in drawing conclusions

**Responsibility towards resources and environments**

- Show a sense of personal responsibility for maintaining a sustainable environment
- Demonstrate awareness of the environmental consequences of individual actions
- Demonstrate willingness to take action to maintain natural resources

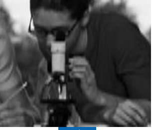
would serve as stimulus material, the competencies required to respond to the questions or issues, or the knowledge and attitudes central to the exercise.

A test unit comprises specific stimulus material, which may be a brief written passage, or text accompanying a table, chart, graph, or diagram, plus items which are a set of independently scored questions of various types, as illustrated by the three examples (*CATCHING THE KILLER, MALARIA* and *SCHOOL MILK STUDY*) already discussed and the additional examples included in the Annex A.

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

PISA 2006 test units incorporate up to four cognitive items which assess students' scientific competencies. Each item involves the predominant use of one of the scientific competencies and requires mainly knowledge of science or knowledge about science. In most cases, more than one competency and more than one knowledge category was assessed (by different items) within a unit.

Four types of items were used to assess the competencies and scientific knowledge identified in the framework. About one-third of the items were (simple) multiple-choice items, which required the



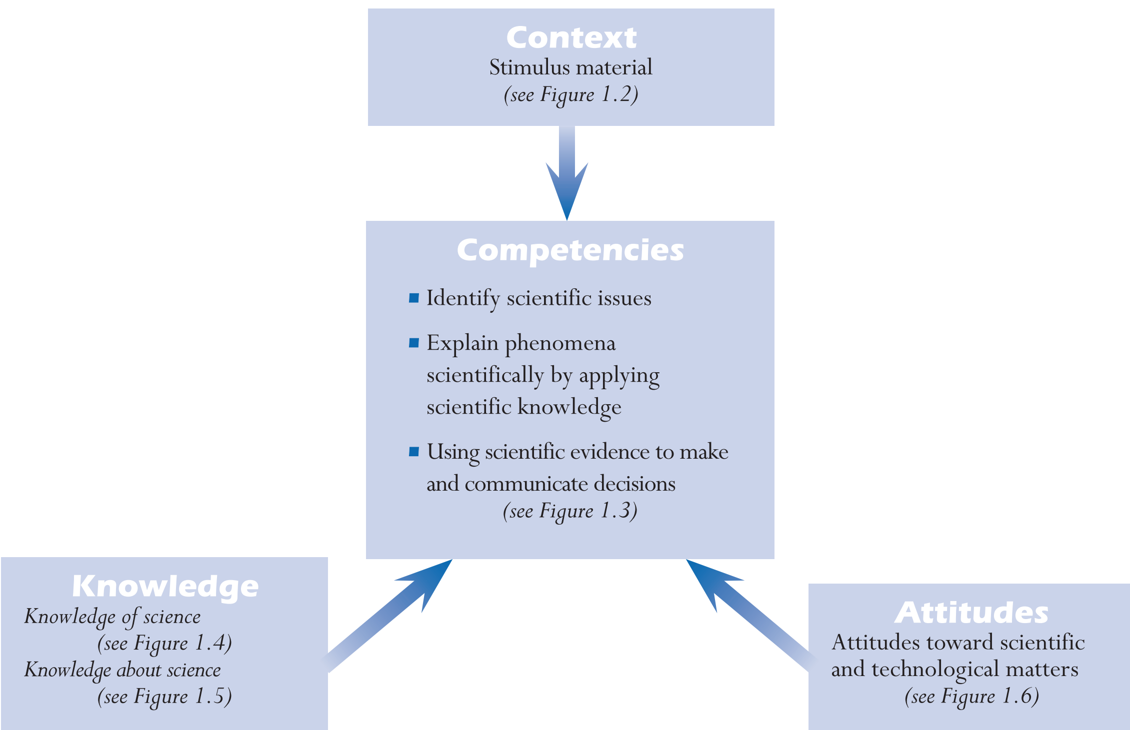


selection of a single response from four options. A further third either required short constructed responses, like Question 1 of *MALARIA* (Science Example 2), or complex multiple-choice items. Question 1 of *SCHOOL MILK STUDY* (Science Example 3), which requires students to respond to a series of related "Yes/No" questions, is a typical complex multiple-choice item. The remaining one-third of the items were open-constructed response items, like Question 2 in *SCHOOL MILK STUDY* (Science Example 3), that required a relatively extended written or drawn response from a student.

Multiple-choice and short-constructed response items can be used to validly assess most of the cognitive processes involved in the three scientific competencies and open-response items provide the opportunity to assess the ability to communicate.

Although the majority of the items are dichotomously scored (that is, credit or no credit), some of the complex multiple-choice and open-response items will involve partial credit scoring, which give students credit for getting part of the question correct, but not the whole question. For each partial credit item, a detailed coding guide that allows for "Full credit", "Partial credit" and "No credit" is provided. The categories "Full credit", "Partial credit" and "No credit" divide students' responses into three groups in terms of the extent to which the students demonstrate ability to answer the question. A "Full credit" response, although not necessarily absolutely scientifically correct, requires a student to exhibit a level of understanding of the topic appropriate for a scientifically literate 15-year-old student. Less sophisticated or correct responses may qualify for "Partial credit", with completely incorrect, irrelevant or missing responses being assigned "No credit". Question 1 of *MALARIA* (Science Example 2) is a partial credit item and its scoring scheme (coding guide) is shown in Science Example 4.

Figure 1.7 ■ **A tool for constructing and analysing assessment units and items**







## Science Example 4: MALARIA (Question 1 scoring)

**Full Credit**

Code 2: All three correct: [1 and 3]; [2]; and [1, 3 and 4] in that order

**Partial Credit**

Code 1: Two of the three rows correct

OR

One (or more) correct, **but none wrong**, in each row

**No Credit**

Code 0: Other responses

Code 9: Missing

---

Most of the new units included in the PISA 2006 science test also contain an item that assesses students' *interest in learning about science* or an item that assesses *support for scientific enquiry* or both types of items. Question 3 of the unit *CATCHING THE KILLER*, included below as Science Example 5, is an example of this. This item requires students to indicate their level of interest in three tasks to assess their interest in learning more about the application of science to solving crime. A unipolar response format ("High interest", "Medium interest", "Low interest", "No interest"), rather than the conventional bipolar one ("Strongly agree", "Agree", "Disagree", "Strongly disagree"), is used in this example to reduce the influence of social desirability on responses.

## Science Example 5: CATCHING THE KILLER (Attitudinal item)

### Question 3: CATCHING THE KILLER

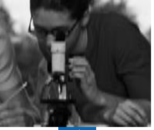
How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing more about the use of DNA in solving crime. | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| b) | Learning more about how DNA profiling works. | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| c) | Understanding better how crime can be solved using science. | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |

---

On the actual test form given to students, attitudinal items are distinctively formatted in a shaded box to remind students that, for each statement, they should tick the box that indicates their own opinion about the statement. In addition, the general directions at the start of each booklet include the following instruction:





Some of the questions are about your attitude or opinion regarding certain issues. These questions are set out differently from the others – they appear inside a shaded box. THERE IS NO CORRECT ANSWER to these questions and they will not count in your test score, but it is important that you answer them truthfully.

The need for students to have a degree of *reading literacy* in order to understand and answer written questions on *scientific literacy* raises an issue of the level of *reading literacy* required. Stimulus material and questions used language that is as clear, simple and brief as possible while still conveying the appropriate meaning. The number of concepts introduced per paragraph was limited and questions that predominantly assess *reading literacy* or *mathematical literacy* were avoided.

## Science assessment structure

It is important that the test includes an appropriate balance of items assessing the various components of the scientific literacy framework. Figure 1.8 shows the desired balance of items relating to the knowledge of science versus knowledge about science. The balance is expressed in terms of the percentage of total score points allocated to each category. Figure 1.8 also shows the desired distribution of score points among the various knowledge of science and knowledge about science categories.

Figure 1.8 ■ **Desired distribution of score points for knowledge**

| ***Knowledge of science*** | **Per cent of score points** |
|---|---|
| Physical systems | 15-20 |
| Living systems | 20-25 |
| Earth and space systems | 10-25 |
| Technological systems | 5-10 |
| *Subtotal* | *60-65* |
| ***Knowledge about science*** | |
| Scientific enquiry | 15-20 |
| Scientific explanation | 15-20 |
| *Subtotal* | *35-40* |
| ***Total*** | ***100*** |

The desired balance for scientific competencies is given in Figure 1.9.

Figure 1.9 ■ **Desired distribution of score points for scientific competencies**

| ***Scientific competencies*** | **Per cent of score points** |
|---|---|
| Identifying scientific issues | 25-30 |
| Explaining phenomenas scientifically | 35-40 |
| Using scientific evidence | 35-40 |
| ***Total*** | ***100*** |



Item contexts are spread across personal, social and global settings roughly in the ratio 1:2:1. A wide selection of areas of application were used for units, subject to satisfying as far as possible the various constraints imposed by the previous two paragraphs.

About 60% of the units contain one or two attitudinal items that assess students' *interest in learning about science* or their *support for scientific enquiry*. Responding to these items occupies about 11% of the total test time. To facilitate comparability of performance over time, link items included from the two previous PISA science assessments did not contain attitudinal items.

## Reporting scales

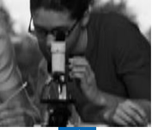
To meet the aims of PISA, the development of scales of student achievement is essential. The process of arriving at a scale has to be iterative. Initial descriptions, based on the results of the trials and the PISA 2000 and 2003 surveys – and informed by past experience of assessing science achievement and findings from research into learning and cognitive development in science – are likely to be modified as more data are accumulated in this and future surveys.

The construction of scales is facilitated by the inclusion of items which have a wide range of difficulties. Factors that determine difficulty in items assessing science achievement include the:

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

For PISA 2000, when science was a minor domain and thus having limited testing time, students' science achievement was reported in terms of a proficiency scale with a mean of 500 and a standard deviation of 100. Although no proficiency levels were identified, it was possible to describe what processes (*i.e.* scientific competencies) students can perform at three points in this scale (OECD, 2001):

- Towards the top end of the scientific literacy scale (around 690 points) students are generally able to create or use conceptual models to make predictions or give explanations; to analyse scientific investigations in order to grasp, for example, the design of an experiment or to identify an idea being tested; to compare data in order to evaluate alternative viewpoints or differing perspectives; and to communicate scientific arguments and/or descriptions in detail and with precision.
- At around 550 points, students are typically able to use scientific knowledge to make predictions or provide explanations; to recognise questions that can be answered by scientific investigation and/or identify details of what is involved in a scientific investigation; and to select relevant information from competing data or chains of reasoning in drawing or evaluating conclusions.





- Towards the lower end of the scale (around 400 points), students are able to recall simple factual scientific knowledge (*e.g.* names, facts, terminology, simple rules); and to use common scientific knowledge in drawing or evaluating conclusions.

For PISA 2003, the reporting of science results followed a similar format to that of 2000 (OECD, 2004). However, with science as the major domain for the PISA 2006 assessment, the increased testing time available should enable the construction of separate scales based on either the scientific competencies or the two knowledge components.

Proficiency in science in PISA 2000 and 2003 was described on a scale in terms of scientific competencies as shown in the Figure 1.3. By examining the descriptions we can derive the skeleton of each PISA 2006 competency scale. For example, the skeleton scale shown in Figure 1.10 can be derived from the competency "Using scientific evidence".

Figure 1.10 ■ **Example of a competency-based reporting scale**

| | |
|---|---|
| High | Can compare data in order to evaluate alternative viewpoints or differing perspectives; can communicate scientific arguments and/or descriptions in detail and with precision. |
| | Able to select relevant information from competing data or chains of reasoning in drawing or evaluating conclusions. |
| Low | Able to use common scientific knowledge in drawing or evaluating conclusions. |

Alternatively, it should be possible to report separate scales for the two knowledge components, knowledge of science and knowledge about science. The competencies would then be central to describing the proficiency levels for these two knowledge scales. Decisions about the actual scales to be reported, and the number of proficiency levels to be identified, will be made following analysis of the PISA 2006 assessment data.

It should also be possible to prepare reliable scales for *interest in science* and *support for scientific enquiry* using the data obtained from the embedded attitudinal items and the student questionnaire. A scale for *responsibility towards resources and environments* will be constructed from data obtained from the student questionnaire.

The scores on attitude items will not be included in an index (or overall score) of *scientific literacy;* rather, they will form a component of a profile of student *scientific literacy*.

## SUMMARY

Science was the major testing domain for the first time in PISA 2006. The definition of *scientific literacy* has been elaborated and expanded from that used in PISA 2000 and 2003. A major innovation is to include students' attitudinal responses towards scientific issues, not just in an accompanying questionnaire but in additional questions about attitudes to scientific issues juxtaposed with test questions relating to the same issues. In addition, there is an increased emphasis on students' understanding of the nature and methodology of science itself (their knowledge about science), and of the role of science-based technology.



The PISA 2006 definition of *scientific literacy* has its origin in the consideration of what 15-year-old students should know, value and be able to do as preparedness for life in modern society. Central to the definition, and the assessment of *scientific literacy,* are the competencies that are characteristic of science and scientific enquiry. The ability of students to perform these competencies depends on their scientific knowledge, both knowledge of the natural world and knowledge about science itself, and their attitudes towards science-related issues.

This framework describes and illustrates the scientific competencies, knowledge and attitudes being assessed in PISA 2006 (see Figure 1.11), and the contexts for test items. Test items were grouped into units with each unit beginning with stimulus material that establishes the context for its items. A combination of item types was used and some items involved partial credit scoring. Attitudinal items were embedded in over half of the units and occupied about 11% of testing time.

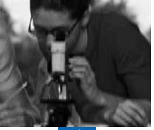
Figure 1.11 ■ **Major components of the PISA 2006 assessment of scientific literacy**

| Competencies | Knowledge | Attitudes |
|---|---|---|
| Identifying scientific issues<br>Explaining scientific phenomena<br>Using scientific eveidence | Knowledge of science:<br>Physical systems<br>Living systems<br>Earth and space systems<br>Knowledge about science:<br>Scientific enquiry<br>Scientific explanations | Interest in science[1]<br>Support for scientific enquiry<br>Responsibility towards resources and environment[2] |

1. Embedded items assess "Interest in learning about science"
2. Not assessed with embedded items

The ratio of items assessing students' knowledge of science, to items assessing their knowledge about science, was about 3:2, while each of the three scientific competencies were assessed by at least 25% of the items. This should enable separate scales, with described proficiency levels, to be constructed for each of the competencies, or for the two types of knowledge. Scales should also be able to be constructed for the attitudes that were assessed with embedded items.

Further examples to illustrate the PISA science assessment framework are included in Annex A.



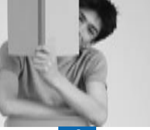


*Notes*

1. Throughout this framework, "natural world" includes the changes made by human activity, including the "material world" designed and shaped by technologies.
2. The PISA science concept of "literacy" can be compared to the DeSeCo (OECD, 2003b) definition of "competency" in that both involve attitudes and values, as well as knowledge and skills.
3. Knowledge of the design or internal working of technology artefacts (*e.g.* aeroplanes, engines, computers) will not be assumed.

2

# Reading Literacy





## DEFINITION OF THE DOMAIN

Definitions of reading and *reading literacy* have changed over time in parallel with changes in society, the economy and culture. The concept of learning, and particularly the concept of lifelong learning, has expanded perceptions of *reading literacy* and the demands made on it. Literacy is no longer considered an ability only acquired in childhood during the early years of schooling. Instead, it is viewed as an expanding set of knowledge, skills and strategies which individuals build on throughout life in various situations and through interaction with their peers and with the larger communities in which they participate.

Through a consensus-building process involving the reading experts selected by the participating countries and the PISA advisory groups, the following definition of *reading literacy* was adopted for the survey:

> *Reading literacy* is understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential and to participate in society.

This definition goes beyond the notion of *reading literacy* as decoding and literal comprehension: it implies that *reading literacy* involves understanding, using and reflecting on written information for a variety of purposes. It thus takes into account the active and interactive role of the reader in gaining meaning from written texts. The definition also recognises the full scope of situations in which *reading literacy* plays a role for young adults, from private to public, from school to work, from active citizenship to lifelong learning. It spells out the idea that literacy enables the fulfilment of individual aspirations – from defined aspirations such as gaining an educational qualification or obtaining a job to those less immediate goals which enrich and extend one's personal life. Literacy also provides the reader with a set of linguistic tools that are increasingly important for meeting the demands of modern societies with their formal institutions, large bureaucracies and complex legal systems.

Readers respond to a given text in a variety of ways as they seek to use and understand what they are reading. This dynamic process involves many factors, some of which can be manipulated in large-scale assessments such as PISA. These include the reading situation, the structure of the text itself and the characteristics of the questions that are asked about the text (the test rubric). All of these factors are regarded as important components of the reading process and were considered in the creation of the items used in the assessment.

In order to use text format, characteristics of the items and situations in constructing the assessment tasks, and later in interpreting the results, the range for each of these factors had to be specified. This allowed for the categorisation of each task so that the weighting of each component could be taken into account in the final assembly of the survey.

## TEXT FORMAT

At the heart of the PISA assessment is a distinction between continuous and non-continuous texts.

- *Continuous texts* are typically composed of sentences that are, in turn, organised into paragraphs. These may fit into even larger structures such as sections, chapters and books. The primary classification of continuous texts is by rhetorical purpose, that is, by text type.

- *Non-continuous texts* (or documents, as they are known in some approaches) can be categorised in two ways. One is the formal structure approach used in the work of Kirsch and Mosenthal (1989-1991). Their work classifies non-continuous texts by the way underlying lists are put. This approach is useful for understanding the similarities and differences between types of non-continuous texts. The other method of classification is by everyday descriptions of the formats of these texts. This second approach is used in classifying non-continuous texts in PISA.

### Continuous texts

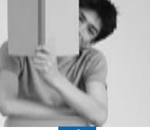
Text types are standard ways of organising continuous texts by content and author's purpose.

- *Narration* is the type of text in which the information refers to properties of objects in time. Narrative texts typically provide answers to "when", or "in what sequence" questions.
- *Exposition* is a type of text in which the information is presented as composite concepts or mental constructs, or elements into which concepts or mental constructs can be analysed. The text provides an explanation of how the component elements interrelate in a meaningful whole and often answers "how" questions.
- *Description* is a type of text in which the information refers to properties of objects in space. Descriptive texts typically provide an answer to "what" questions.
- *Argumentation* is a type of text that presents propositions as to the relationship between concepts, or other propositions. Argumentative texts often answer "why" questions. Another important sub-classification of argumentative texts is persuasive texts.
- *Instruction* (sometimes called injunction) is the type of text that provides directions on what to do and includes procedures, rules, regulations and statutes specifying certain behaviours.
- *Documents or records* are texts designed to standardise and conserve information. They can be characterised by highly formalised textual and formatting features.
- *Hypertext* is a set of text slots linked together in such a way that the units can be read in different sequences, allowing readers to follow various routes to the information.

### Non-continuous texts

*Non-continuous texts* are organised differently from continuous texts and so require different kinds of reading approaches. Classifying non-continuous texts by their format, as shown below, provides a familiar means of discussing what types of non-continuous texts may be included in the assessment.

- *Charts and graphs* are iconic representations of data. They are used for the purposes of scientific argumentation, and also in journals and newspapers to display numerical and tabular public information in a visual format.
- *Tables* are row and column matrices. Typically, all the entries in each column and each row share properties and thus the column and row labels are part of the information structure of the text. Common tables include schedules, spreadsheets, order forms and indexes.
- *Diagrams* often accompany technical descriptions (*e.g.* demonstrating parts of a household appliance), expository texts and instructive texts (*e.g.* illustrating how to assemble a household appliance). It is often useful to distinguish procedural (how to) from process (how something works) diagrams.





- *Maps* are non-continuous texts that indicate the geographical relationships between places. There is a variety of types of maps. Road maps mark the distance and routes between identified places. Thematic maps indicate the relationships between locations and social or physical features.
- *Forms* are structured and formatted texts which request the reader to respond to specific questions in specified ways. Forms are used by many organisations to collect data. They often contain structured or pre-coded answer formats. Typical examples are tax forms, immigration forms, visa forms, application forms, statistical questionnaires, etc.
- *Information sheets* differ from forms in that they provide, rather than request, information. They summarise information in a structured way and in such a format that the reader can easily and quickly locate specific pieces of information. Information sheets may contain various text forms as well as lists, tables, figures and sophisticated text-based graphics (headings, fonts, indentation, borders, etc.) to summarise and highlight information. Timetables, price lists, catalogues and programmes are examples of this type of non-continuous text.
- *Calls and advertisements* are documents designed to invite the reader to do something, *e.g.* to buy goods or services, attend gatherings or meetings, elect a person to a public office, etc. The purpose of these documents is to persuade the reader. They offer something and request both attention and action. Advertisements, invitations, summonses, warnings and notices are examples of this document format.
- *Vouchers* testify that their owner is entitled to certain services. The information that they contain must be sufficient to show whether the voucher is valid or not. Typical examples are tickets, invoices, etc.
- *Certificates* are written acknowledgements of the validity of an agreement or a contract. They are formalised in content rather than format. They require the signature of one or more persons authorised and competent to bear testimony of the truth of the given statement. Warranties, school certificates, diplomas, contracts, etc. are documents that have these properties.

The distribution and variety of texts that students are asked to read for PISA are important characteristics of the assessment. Figure 2.1 shows the distributions of tasks for continuous and non-continuous texts in PISA 2000 (reading as major domain) and in PISA 2003 and 2006 (reading as minor domain). It can be readily seen that in 2000, 2003 and 2006 cycles continuous texts represent two-thirds of the tasks or items contained in the assessment. Within this category, in the three cycles, the largest percentage comes from expository texts.

## CHARACTERISTICS OF THE ITEMS

Three sets of variables are used to describe the characteristics of the items: the processes (aspects), which set out the task for the student; item types, which set out the ways in which students are asked to demonstrate their proficiency at the task; and rules for coding, which specify how students' answers are to be evaluated. Each of these will be discussed in turn, though the first requires considerably more attention.

### Five processes (aspects)

In an effort to simulate authentic reading situations, the PISA reading assessment measures the following five processes associated with achieving a full understanding of a text, whether the text

Figure 2.1 ■ **Distribution of reading literacy tasks, by text format and type**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003 and 2006)

| Text format and type | Percentage of tasks by text format and type (%) | | Percentage of tasks by text format and type, based on the whole test (%) | |
|---|---|---|---|---|
| | Reading as a major domain (PISA 2000) | Reading as a minor domain (PISA 2003 and 2006) | Reading as a major domain (PISA 2000) | Reading as a minor domain (PISA 2003 and 2006) |
| ***Continuous*** | | | | |
| Narrative | 21 | 17 | 14 | 11 |
| Expository | 36 | 67 | 24 | 43 |
| Descriptive | 14 | 17 | 9 | 11 |
| Argumentative and persuasive | 20 | - | 13 | - |
| Injunctive | 10 | - | 7 | - |
| **Total[1]** | **100** | **100** | **68** | **64** |
| ***Non-continuous*** | | | | |
| Charts and graphs | 37 | 20 | 12 | 7 |
| Tables | 29 | 40 | 9 | 14 |
| Diagrams | 12 | - | 4 | - |
| Maps | 10 | 10 | 3 | 4 |
| Forms | 10 | 30 | 3 | 11 |
| Advertisements | 2 | - | 1 | - |
| **Total[1]** | **100** | **100** | **34** | **37** |

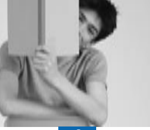
1. Data may not always add up to the totals indicated because of roundings.

is continuous or non-continuous. Students are expected to demonstrate their proficiency in all of these processes:

- *Retrieving information*
- *Forming a broad general understanding*
- *Developing an interpretation*
- *Reflecting on and evaluating the content of a text*
- *Reflecting on and evaluating the form of a text*

The full understanding of texts involves all of these processes. It is expected that all readers, irrespective of their overall proficiency, will be able to demonstrate some level of competency in each of them (Langer, 1995). While there is an interrelationship between the five aspects – each may require many of the same underlying skills – successfully accomplishing one may not ensure successful completion of any other. Some view them as being in the repertoire of each reader at every developmental level rather than forming a sequential hierarchy or set of skills.

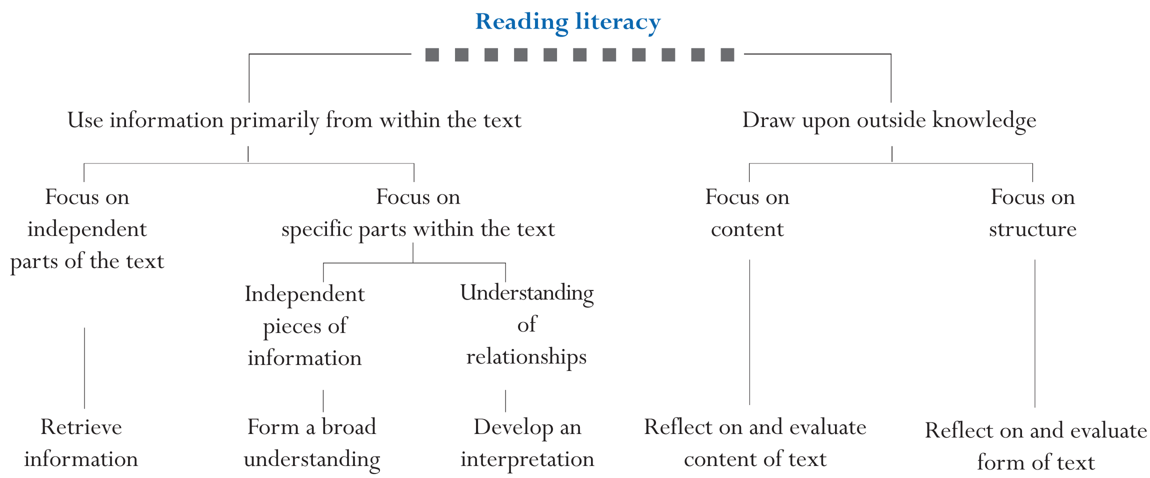
Figure 2.2 identifies the key distinguishing characteristics of the five processes of reading measured in PISA. While this figure necessarily oversimplifies each process, it provides a useful scheme for organising and remembering the relationships between them. As depicted in this figure, the five processes can be distinguished in terms of four characteristics. The first deals with the extent to which the reader is expected to use information primarily from within the text or to draw also upon outside knowledge. A second characteristic involves the extent to which the reader is asked to focus on independent parts of the text or on the relationships within the information contained





in the text. Sometimes readers are expected to retrieve independent pieces of information while at other times they are asked to demonstrate their understanding of the relationships between parts of the text. Focusing on either the whole text or on relationships between parts of the text is the third distinguishing characteristic. The fourth characteristic relates to whether the reader is asked to deal with the content or substance of the text rather than its form or structure. The five processes of reading are represented in the last line of Figure 2.2 at the ends of the various branches. By starting at the top of the figure and following each branch one can see which characteristics are associated with each process.

Figure 2.2 ■ **Characteristics distinguishing the five processes (aspects) of *reading literacy***




The following discussion attempts to define each process operationally and to associate it with particular kinds of items. Although each process is discussed in terms of a single text, each can also apply to multiple texts when these are presented together as a unit within the test. The description of each process has two parts. The first provides a general overview of the process, while the second describes particular ways in which the process might be assessed.

## Retrieving information

In the course of daily life, readers often need a particular piece of information: a telephone number or the departure time for a bus or train. They may want to find a particular fact to support or refute a claim someone has made. In situations such as these, readers are interested in retrieving isolated pieces of information. To do so, readers must scan, search for, locate and select relevant information. The processing involved is most frequently at the sentence level, though in some cases the information may be in two or more sentences or in different paragraphs.

In assessment tasks that call for retrieving information, students must match information given in the question with either identically worded or synonymous information in the text and use this to find the new information called for. In these tasks, *retrieving information* is based on the text itself and on explicit information included in it. Retrieving tasks require the student to find information based on requirements or features specified in questions. The student has to detect or identify one



or more essential elements of a question: characters, place/time, setting, etc. and then to search for a match that may be literal or synonymous.

Retrieving tasks can involve various degrees of ambiguity. For example, the student may be required to select explicit information, such as an indication of time or place in a text or table. A more difficult version of this same type of task might involve finding synonymous information. This sometimes involves categorisation skills, or it may require discriminating between two similar pieces of information. The different levels of proficiency associated with this process of comprehension can be measured by systematically varying the elements that contribute to the difficulty of the task.

### Forming a broad general understanding

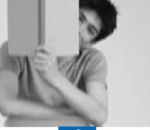
To *form a broad general understanding* of what has been read, a reader must consider the text as a whole or in a broad perspective. There are various assessment tasks in which readers are asked to form a broad general understanding. Students may demonstrate initial understanding by identifying the main topic or message or by identifying the general purpose or use of the text. Examples include tasks that require the reader to select or create a title or thesis for the text, to explain the order of simple instructions, or to identify the main dimensions of a graph or a table. Others include tasks that require the student to describe the main character, setting or milieu of a story, to identify a theme or message of a literary text, or to explain the purpose or use of a map or a figure.

Within this process some tasks might require the student to match a particular piece of text to the question. For example, this would happen when a theme or main idea is explicitly stated in the text. Other tasks may require the student to focus on more than one specific reference in the text – for instance, if the reader had to deduce the theme from the repetition of a particular category of information. Selecting the main idea implies establishing a hierarchy among ideas and choosing the most general and overarching. Such a task indicates whether the student can distinguish between key ideas and minor details, or can recognise the summary of the main theme in a sentence or title.

### Developing an interpretation

*Developing an interpretation* requires readers to extend their initial impressions so that they develop a more specific or complete understanding of what they have read. Tasks in this category call for logical understanding; readers must process the organisation of information in the text. To do so, readers must demonstrate their understanding of cohesion even if they cannot explicitly state what cohesion is. In some instances, developing an interpretation may require the reader to process a sequence of just two sentences relying on local cohesion, which might even be facilitated by the presence of cohesive markers, such as the use of "first" and "second" to indicate a sequence. In more difficult instances (*e.g.* to indicate relations of cause and effect), there might not be any explicit markings.

Examples of tasks that might be used to assess this process include comparing and contrasting information, drawing inferences, and identifying and listing supporting evidence. "Compare and contrast" tasks require the student to draw together two or more pieces of information from the text. In order to process either explicit or implicit information from one or more sources in such tasks, the reader must often infer an intended relationship or category. This process of comprehension is also assessed in tasks that require the student to make inferences about the author's intention, and to identify the evidence used to infer that intention.





### Reflecting on and evaluating the content of a text

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

Assessment tasks representative of this category of processing include providing evidence or arguments from outside the text, assessing the relevance of particular pieces of information or evidence, or drawing comparisons with moral or aesthetic rules (standards). The student might be asked to offer or identify alternative pieces of information that might strengthen an author's argument, or to evaluate the sufficiency of the evidence or information provided in the text.

The outside knowledge to which textual information is to be connected may come from the student's own knowledge, from other texts provided in the assessment, or from ideas explicitly provided in the question.

### Reflecting on and evaluating the form of a text

Tasks in this category require readers to stand apart from the text, consider it objectively and evaluate its quality and appropriateness. Knowledge of such things as text structure, genre and register play an important role in these tasks. These features, which form the basis of an author's craft, figure strongly in understanding standards inherent in tasks of this nature. Evaluating how successful an author is in portraying some characteristic or persuading a reader depends not only on substantive knowledge but also on the ability to detect nuances in language – for example, understanding when the choice of an adjective might colour interpretation.

Some examples of assessment tasks characteristic of *reflecting on and evaluating the form of a text* include determining the utility of a particular text for a specified purpose and evaluating an author's use of particular textual features in accomplishing a particular goal. The student may also be called upon to describe or comment on the author's use of style and to identify the author's purpose and attitude.

### Distribution of tasks

Figure 2.3 shows the distribution of reading literacy tasks by each of the three subscales generated from the five reading processes (aspects) defined above. The largest category of tasks, which accounts for approximately 50% of the test, is represented by the two branches of Figure 2.2 that ask students to focus on relationships within a text. These tasks require students either to form a broad understanding or to develop an interpretation. They have been grouped together for reporting purposes into a single process called Interpreting texts. In PISA 2000, 2003 and 2006, the next largest category was made up of the 29% of the tasks that require students to demonstrate their skill at retrieving isolated pieces of information. Each of these processes – forming a broad understanding, retrieving information and developing an interpretation – focuses on the degree



Figure 2.3 ■ **Distribution of reading literacy tasks, by reading process (aspect)**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003 and 2006)

| Reading process (aspect) | Percentage of tasks | |
|---|---|---|
| Retrieving information | 29 | 29 |
| Interpreting texts | 49 | 50 |
| Reflection and evaluation | 22 | 21 |
| **Total[1]** | **100** | **100** |

1. Data may not always add up to the totals indicated because of rounding.

to which the reader can understand and use information contained primarily within the text. The remaining tasks, approximately 20%, required students to reflect on either the content or information provided in the text or on the structure and form of the text itself.

## Item types

The reading tasks in PISA are made up of various types, including multiple choice and various constructed tasks which require the students to write their answers rather than simply select from a number of given responses. The different types of tasks also require different marking. Figure 2.4 indicates that in PISA 2000, 2003 and 2006, around 43% of the reading literacy tasks in the PISA assessment were open-constructed response items which required judgement on the part of the marker. The remaining tasks consist of closed constructed-response items that require little judgement on the part of the marker, as well as simple multiple-choice items, for which students choose one of several alternative answers, and complex multiple-choice items, for which students choose more than one response.

This table also reveals that while multiple choice and open-constructed response items are represented across the processes, they are not distributed evenly.

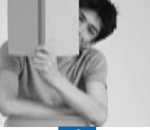
Figure 2.4 ■ **Distribution of reading literacy tasks, by reading process (aspect) and item type**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003 and 2006)

| | Item types | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Process (aspect) | Percentage of multiple-choice items | | Percentage of complex multiple-choice items | | Percentage of closed-constructed response items | | Percentage of open-constructed response items[1] | | Total[2] | |
| Retrieving information | 8 | - | 2 | 4 | 6 | 14 | 13 | 11 | **29** | **29** |
| Interpreting texts | 32 | 29 | 2 | 4 | 2 | 7 | 13 | 11 | **49** | **50** |
| Reflection and evaluation | 2 | - | 2 | - | - | - | 18 | 21 | **22** | **21** |
| **Total[2]** | **42** | **29** | **6** | **7** | **9** | **21** | **44** | **43** | **100** | **100** |

1. This category includes short-response items.
2. Data may not always add up to the total indicated because of rounding.





A larger percentage of multiple-choice items are associated with the two processes dealing with interpreting relationships within a text. This is shown in the second row of Figure 2.4. In contrast, while reflection and evaluation tasks account for around 20% in PISA 2000, 2003 and 2006, only 2% in 2000 are multiple choice. Of the reflection and evaluation tasks, around 20% are open-constructed response items that require judgement on the part of the marker.

### Marking

Marking is relatively simple with dichotomously scored multiple-choice items: the student has either chosen the designated answer or not. Partial-credit models allow for more complex marking of items. Here, because some "wrong" answers are more close to being correct than others, students who provide an "almost right" answer receive partial credit. Psychometric models for such polytomous scoring are well-established and in some ways are preferable to dichotomous scoring, as they utilise more of the information in the responses. Interpretation of polytomous marking is more complex, however, as each task has several locations on the difficulty scale: one for the full-credit answer and others for each of the partial-credit answers. Partial-credit marking is used for some of the more complex constructed-response items in PISA.

## SITUATIONS

The manner in which situation was defined was borrowed from the Council of Europe's work on language. Four situation variables were identified: *reading for private use, reading for public use, reading for work* and *reading for education*. While the intention of the PISA reading literacy assessment was to measure the kinds of reading that occur both within and outside classrooms, the manner in which situation was defined could not be based simply on where the reading activity is carried out. For example, textbooks are read both in schools and in homes, and the process and purpose of reading these texts differ little from one setting to another. Moreover, reading also involves the author's intended use, different types of content and the fact that others (*e.g.* teachers and employers) sometimes decide what should be read and for what purpose.

Thus, for the purpose of this assessment, situation can be understood as a general categorisation of texts based on the author's intended use, on the relationship with other persons implicitly or explicitly associated with the text, and on the general content. The sample texts were drawn from a variety of situations to maximise the diversity of content included in the reading literacy survey. Close attention was also paid to the origin of texts selected for inclusion in this survey. The goal was to reach a balance between the broad definition of *reading literacy* used in PISA and the linguistic and cultural diversity of participating countries. This diversity helped to ensure that no one group would be either advantaged or disadvantaged by the assessment content.

The four situation variables taken from the work of the Council of Europe are described as follows:

- *Reading for private use (personal):* This type of reading is carried out to satisfy an individual's own interests, both practical and intellectual. It also includes reading to maintain or develop personal connections to other people. Contents typically include personal letters, fiction, biography and informational texts read for curiosity, as a part of leisure or recreational activities.



- *Reading for public use:* This type of reading is carried out to participate in the activities of the wider society. It includes the use of official documents as well as information about public events. In general, these tasks are associated with more or less anonymous contact with others.
- *Reading for work (occupational):* While not all 15-year-olds will actually have to read at work, it is important to assess their readiness to move into the world of work since, in most countries, over 50% of them will be in the labour force within one to two years. The prototypical tasks of this type are often referred to as "reading to do" (Sticht, 1975; Stiggins, 1982) in that they are tied to the accomplishment of some immediate task.
- *Reading for education:* This type of reading is normally involved with acquiring information as part of a larger learning task. The materials are often not chosen by the reader, but assigned by a teacher. The content is usually designed specifically for the purpose of instruction. The prototypical tasks are those usually identified as "reading to learn" (Sticht, 1975; Stiggins, 1982).

Figure 2.5 shows the distribution of reading literacy tasks in the assessment across all four situations for two scenarios: when reading was a major domain (PISA 2000) and when it is a minor domain (PISA 2003 and 2006).

Figure 2.5 ■ **Distribution of reading literacy tasks, by situation**

Reading as a major domain (PISA 2000)

Reading as a minor domain (PISA 2003 and 2006)

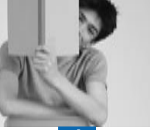
| Situation | Percentage of tasks | |
|---|---|---|
| | Reading as a major domain (PISA 2000) | Reading as a minor domain (PISA 2003 and 2006) |
| Personal | 20 | 21 |
| Public | 38 | 25 |
| Occupational | 14 | 25 |
| Educational | 28 | 29 |
| **Total** | **100** | **100** |

## REPORTING OUTCOMES

### Scaling the reading literacy tasks

The reading literacy tasks are constructed and administered to nationally representative samples of 15-year-old students in participating countries to ensure that the assessment provides the broadest possible coverage of *reading literacy* as defined here. However, no individual student can be expected to respond to the entire set of tasks. Accordingly, the survey is designed to give each student participating in the study a subset of the total pool of tasks, while at the same time ensuring that each of the tasks is administered to nationally representative samples of students. Summarising the performance of students across this entire pool of tasks thus poses a challenge.

One may imagine the reading literacy tasks arranged along a continuum in terms of difficulty for students and the level of skill required to answer each item correctly. The procedure used in PISA to capture this continuum of difficulty and ability is Item Response Theory (IRT). IRT is a mathematical model used for estimating the probability that a particular person will respond correctly to a given



task from a specified pool of tasks. This probability is modelled along a continuum which summarises both the proficiency of a person in terms of his or her ability and the complexity of an item in terms of its difficulty. This continuum of difficulty and proficiency is referred to as a "scale".

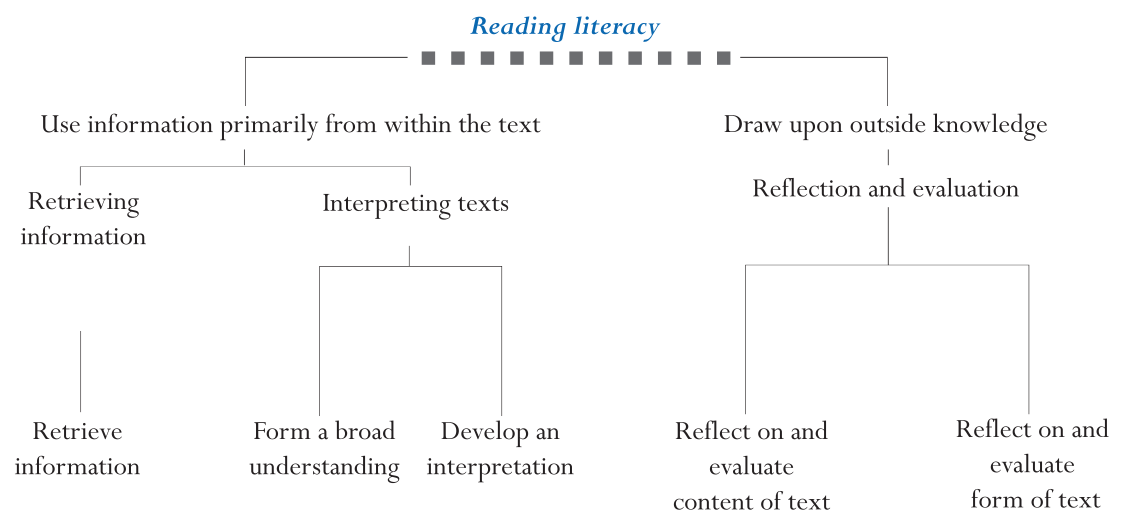
## Reporting

PISA 2006 will follow the reporting scheme used in PISA 2000 and 2003, which reported outcomes in terms of a proficiency scale based on theory and interpretable in policy terms. The results of the reading literacy assessment were first summarised on a single composite reading literacy scale having a mean of 500 and a standard deviation of 100. In addition, student performance was also represented on five subscales: three process (aspect) subscales ("retrieving information", "interpreting texts", and "reflection and evaluation") (OECD, 2001) and two text format subscales (*continuous* and *non-continuous text*) (OECD, 2002). These five subscales make it possible to compare mean scores and distributions among subgroups and countries by various components of the reading literacy construct. Although there is a high correlation between these subscales, reporting results on each subscale may reveal interesting interactions among the participating countries. Where such features occur, they can be examined and linked to the curriculum and teaching methodology used. In some countries, the important question may be how to teach the current curriculum better. In others, the question may not only be how to teach but also what to teach.

### The reading process (aspect) subscales

Figure 2.6 summarises the reading literacy tasks in terms of three processes. There are two reasons for reducing the number of processes from five to three for reporting purposes. The first is pragmatic. In 2003 and 2006, reading, as a minor domain, is restricted to about 30 items instead of the 141 that were used in 2000 when reading was a major domain. The amount of information, therefore, is insufficient to report trends over five process subscales. The second reason is conceptual. The three process subscales are based on the set of five processes shown in Figure 2.2. *Forming a broad understanding* and *developing an interpretation* have been grouped together in an "interpreting texts"

Figure 2.6 ■ **Relationship between the reading literacy framework and the process (aspect) subscales**





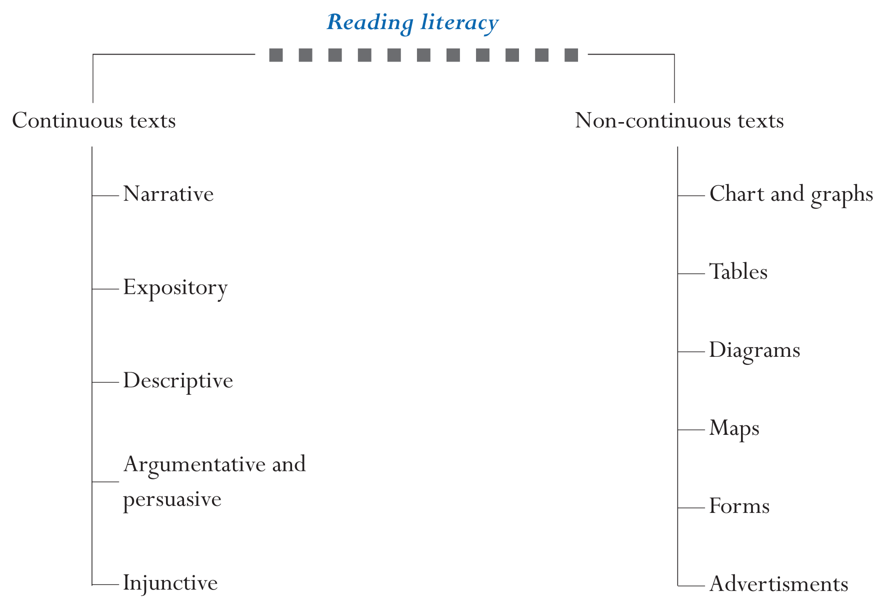
subscale because, in both, the reader processes information in the text: in the case of *forming a broad understanding*, the whole text and in the case of *developing an interpretation*, one part of the text in relation to another. *Reflecting on and evaluating the content of a text* and *reflecting on and evaluating the form of a text* have been collapsed into a single “reflection and evaluation” subscale because the distinction between reflecting on and evaluating content and reflecting on and evaluating form, in practice, was found to be somewhat arbitrary.

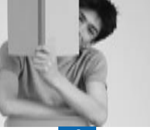
### The text format subscales

PISA 2003 and 2006 also offer the possibility of providing results based on text format subscales, as reported in *Reading for Change: Performance and Engagement across Countries* (OECD, 2002). Figure 2.7 summarises the various text formats and the associated tasks along the two text format subscales. Organising the data in this way provides the opportunity to examine to what extent countries differ with respect to ability to deal with texts in different formats. In reporting results for 2000, two-thirds of the tasks were used to create the continuous text subscale while the remaining one-third of the tasks was used to create the non-continuous text subscale. There is a similar distribution of tasks between the two text formats in PISA 2003 and 2006.

The scores on the composite scale as well as on each of the five subscales represent varying degrees of proficiency. A low score indicates that a student has very limited knowledge and skills, while a high score indicates that a student has quite advanced knowledge and skills. Use of IRT makes it possible not only to summarise results for various subpopulations of students, but also to determine the relative difficulty of the reading literacy tasks included in the survey. In other words, just as individuals receive a specific value on a scale according to their performance in the assessment tasks, each task receives a specific value on a scale according to its difficulty, as determined by the performance of students across the various countries that participate in the assessment.

Figure 2.7 ■ **Relationship between the reading literacy framework and the text format subscales**







## Building an item map

The complete set of reading literacy tasks used in PISA varies widely in text format, situation and task requirements, and hence also in difficulty. This range is captured through what is known as an item map. The item map provides a visual representation of the reading literacy skills demonstrated by students along the scales. The map should contain a brief description of a selected number of released assessment tasks along with their scale values. These descriptions take into consideration the specific skills the item is designed to assess and, in the case of open-ended tasks, the criteria used for judging the item correct. An examination of the descriptions provides some insight into the range of processes required of students and the proficiencies they need to demonstrate at various points along the reading literacy scales.

Figure 2.8 shows an example of an item map from PISA 2000. An explanation of how to interpret it may be useful. The score assigned to each item is based on the theory that someone at a given point on the scale is equally proficient in all tasks at that point on the scale. It was decided that, for the purposes of PISA, "proficiency" should mean that students at a particular point on the reading literacy scale would have a 62% chance of responding correctly to items at that point. For example, in Figure 2.8 an item appears at 421 on the composite scale. This means that students scoring 421 on the composite reading literacy scale will have a 62% chance of correctly answering items graded 421 on the scale. This does not mean that students receiving scores below 421 will always answer incorrectly. Rather, students scoring below 421 will be expected to answer correctly an item of that level of difficulty less than 62% of the time. Conversely, students having scores above 421 will have a greater than 62% chance of responding correctly. It should be noted that the item will also appear on a process subscale and on a format subscale as well as on the combined reading literacy scale. In this example, the item at 421 on the composite scale requires students to identify the purpose that two short texts have in common by comparing the main ideas in each of them. It is an interpretation item and thus appears on the interpreting texts scale as well as on the continuous texts scale.

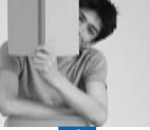
## Levels of reading literacy proficiency

Just as students within each country are sampled to represent the national population of 15-year-old students, each reading literacy task represents a class of tasks from the reading literacy domain. Hence, it represents proficiency in a type of processing and in dealing with a type of text that 15-year-old students should have acquired. One obvious question is, what distinguishes tasks at the lower end of the scale from those in the middle and upper ranges of the scale? Also, do tasks that fall around the same place on the scale share some characteristics that result in their having similar levels of difficulty? Even a cursory review of the item map reveals that tasks at the lower end of each scale differ from those at the higher end. A more careful analysis of the range of tasks along each scale provides indications of an ordered set of information-processing skills and strategies. Members of the reading expert group examined each task to identify a set of variables that seemed to influence its difficulty. They found that difficulty is in part determined by the length, structure and complexity of the text itself. However, they also noted that in most reading units (a unit being a text and a set of questions), the questions range across the reading literacy scale. This means that while the structure of a text contributes to the difficulty of an item, what the reader has to do with that text, as defined by the question or directive, interacts with the text and affects the overall difficulty.



### Figure 2.8 ■ An example of a PISA 2000 item map

| Composite item map | Types of process (Aspect) | | | Text format | |
|---|---|---|---|---|---|
| | Retrieving information | Interpreting | Reflection and evaluation | Continuous | Non-continuous |
| 822 **HYPOTHESISE** about an unexpected phenomenon by taking account of outside knowledge along with all relevant information in a **COMPLEX TABLE** on a relatively unfamiliar topic. (Score 2) | | | ○ | | ○ |
| 727 **ANALYSE** several described cases and **MATCH** to categories given in a **TREE DIAGRAM**, where some of the relevant information is in footnotes. (Score 2) | | ○ | | | ○ |
| 705 **HYPOTHESISE** about an unexpected phenomenon by taking account of outside knowledge along with some relevant information in a **COMPLEX TABLE** on a relatively unfamiliar topic. (Score 1) | | | ○ | | ○ |
| 652 **EVALUATE** the ending of a LONG NARRATIVE in relation to its implicit theme or mood. (Score 2) | | | ○ | ○ | |
| 645 **RELATE NUANCES OF LANGUAGE** in a **LONG NARRATIVE** to the main theme, in the presence of conflicting ideas. (Score 2) | | ○ | | ○ | |
| 631 **LOCATE** information in a **TREE DIAGRAM** using information in a footnote. (Score 2) | ○ | | | | ○ |
| 603 **CONSTRUE** the meaning of a sentence by relating it to broad context in a **LONG NARRATIVE**. | | ○ | | ○ | |
| 600 **HYPOTHESISE** about an authorial decision by relating evidence in a graph to the inferred main theme of **MULTIPLE GRAPHIC DISPLAYS**. | | | ○ | | ○ |
| 581 **COMPARE AND EVALUATE** the style of two open **LETTERS**. | | | ○ | ○ | |
| 567 **EVALUATE** the ending of a **LONG NARRATIVE** in relation to the plot. | | | ○ | ○ | |
| 542 **INFER AN ANALOGICAL RELATIONSHIP** between two phenomena discussed in an open **LETTER**. | | ○ | | ○ | |
| 540 **IDENTIFY** the implied starting date of a **GRAPH**. | ○ | | | | ○ |
| 539 **CONSTRUE THE MEANING** of short quotations from a **LONG NARRATIVE** in relation to atmosphere or immediate situation. (Score 1) | | ○ | | ○ | |
| 537 **CONNECT** evidence from **LONG NARRATIVE** to personal concepts in order to justify opposing points of view. (Score 2) | | | ○ | ○ | |
| 529 **EXPLAIN** a character's motivation by linking events in a **LONG NARRATIVE**. | | ○ | | ○ | |
| 508 **INFER THE RELATIONSHIP** between **TWO GRAPHIC DISPLAYS** with different conventions. | | ○ | | | ○ |
| 486 **EVALUATE** the suitability of a **TREE DIAGRAM** for particular purposes. | | | ○ | | ○ |
| 485 **LOCATE** numerical information in a **TREE DIAGRAM**. | ○ | | | | ○ |
| 480 **CONNECT** evidence from **LONG NARRATIVE** to personal concepts in order to justify a single point of view. (Score 1) | | | ○ | ○ | |
| 478 **LOCATE AND COMBINE** information in a **LINE GRAPH** and its introduction to infer a missing value. | ○ | | | | ○ |
| 477 **UNDERSTAND** the structure of a **TREE DIAGRAM**. | | ○ | | | ○ |
| 473 **MATCH** categories given in a **TREE DIAGRAM** to described cases, when some of the relevant information is in footnotes. | | ○ | | | ○ |
| 447 **INTERPRET** information in a single paragraph to understand the setting of a **NARRATIVE**. | | ○ | | ○ | |
| 445 Distinguish between variables and **STRUCTURAL FEATURES** of a **TREE DIAGRAM**. | | | ○ | | ○ |
| 421 **IDENTIFY** the common **PURPOSE** of **TWO SHORT TEXTS**. | | ○ | | ○ | |
| 405 **LOCATE** pieces of explicit information in a **TEXT** containing strong organizers. | ○ | | | ○ | |
| 397 Infer the **MAIN IDEA** of a simple **BAR GRAPH** from its title. | | ○ | | | ○ |
| 392 **LOCATE** a literal piece of information in a **TEXT** with clear text structure. | ○ | | | ○ | |
| 367 **LOCATE** explicit information in a short, specified section of a **NARRATIVE**. | ○ | | | ○ | |
| 363 **LOCATE** an explicitly stated piece of information in a **TEXT** with headings. | ○ | | | ○ | |
| 356 **RECOGNISE THEME** of an article having a clear subheading and considerable redundancy. | | ○ | | ○ | |





The members of the reading expert group and test developers identified a number of variables that can influence the difficulty of any reading literacy task. One salient factor is the process involved in retrieving information, developing an interpretation or reflecting on what has been read. Processes range in complexity and sophistication from making simple connections between pieces of information, to categorising ideas according to given criteria, and to critically evaluating and hypothesising about a section of text. In addition to the process called for, the difficulty of retrieving information tasks varies with the number of pieces of information to be included in the response, the number of criteria which the information must satisfy, and whether or not what is retrieved needs to be sequenced in a particular way. In the case of interpretative and reflective tasks, the amount of a text that needs to be assimilated is an important factor affecting difficulty. In items that require reflection on the reader's part, difficulty is also conditioned by the familiarity or specificity of the knowledge that must be drawn on from outside the text. In all processes of reading, the difficulty of the task depends on how prominent the required information is, how much competing information is present, and whether or not the reader is explicitly directed to the ideas or information required to complete the task.

In an attempt to capture this progression of complexity and difficulty in PISA 2000, the composite reading literacy scale and each of the subscales were divided into five levels:

| **Level** | **Score points on the PISA scale** |
|---|---|
| 1 | 335 to 407 |
| 2 | 408 to 480 |
| 3 | 481 to 552 |
| 4 | 553 to 625 |
| 5 | More than 625 |

Expert panels judged that the tasks within each level of *reading literacy* shared many of the same task features and requirements, and differed in systematic ways from tasks at higher or lower levels. As a result, these levels appear to be a useful way to explore the progression of reading literacy demands within each scale. This progression is summarised in Figure 2.9. This process was undertaken for *mathematical literacy* in PISA 2003 and will be undertaken for *scientific literacy* in 2006.

### Interpreting the reading literacy levels

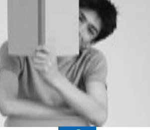
Not only does each level represent a range of tasks and associated knowledge and skills, it also represents a range of proficiencies demonstrated by students. As mentioned previously, the reading literacy levels were initially set by the members of the reading expert group to represent a set of tasks with shared characteristics. These levels also have shared statistical properties. The average student within each level can be expected to successfully perform the average task within that level 62% of the time. In addition, the width of each level is in part determined by the expectation that a student at the lower end of any level will score 50% on any hypothetical test made up of items randomly selected from that level.

Since each reading literacy scale represents a progression of knowledge and skills, students at a particular level not only demonstrate the knowledge and skills associated with that particular level but the proficiencies associated with the lower levels as well. Thus the knowledge and skills assumed



Figure 2.9 ■ **Reading literacy levels map**

<table>
<tr><th></th><th>Retrieving information</th><th>Interpreting texts</th><th>Reflection and evaluation</th></tr>
<tr><td>Level 5</td><td>Locate and possibly sequence or combine multiple pieces of deeply embedded information, some of which may be outside the main body of the text. Infer which information in the text is relevant to the task. Deal with highly plausible and/or extensive competing information.</td><td>Either construe the meaning of nuanced language or demonstrate a full and detailed understanding of a text.</td><td>Critically evaluate or hypothesise, drawing on specialised knowledge. Deal with concepts that are contrary to expectations and draw on a deep understanding of long or complex texts.</td></tr>
<tr><td></td><td colspan="3">Continuous texts: Negotiate texts whose discourse structure is not obvious or clearly marked, in order to discern the relationship of specific parts of the text to its implicit theme or intention.<br>Non-continuous texts: Identify patterns among many pieces of information presented in a display which may be long and detailed, sometimes by referring to information external to the display. The reader may need to realise independently that a full understanding of the section of text requires reference to a separate part of the same document, such as a footnote.</td></tr>
<tr><td>Level 4</td><td>Locate and possibly sequence or combine multiple pieces of embedded information, each of which may need to meet multiple criteria, in a text with unfamiliar context or form. Infer which information in the text is relevant to the task.</td><td>Use a high level of text-based inference to understand and apply categories in an unfamiliar context, and to construe the meaning of a section of text by taking into account the text as a whole. Deal with ambiguities, ideas that are contrary to expectation and ideas that are negatively worded.</td><td>Use formal or public knowledge to hypothesise about or critically evaluate a text. Show accurate understanding of long or complex texts.</td></tr>
<tr><td></td><td colspan="3">Continuous texts: Follow linguistic or thematic links over several paragraphs, often in the absence of clear discourse markers, in order to locate, interpret or evaluate embedded information or to infer psychological or metaphysical meaning.<br>Non-continuous texts: Scan a long, detailed text in order to find relevant information, often with little or no assistance from organisers such as labels or special formatting, and to locate several pieces of information to be compared or combined.</td></tr>
<tr><td>Level 3</td><td>Locate, and in some cases recognise, the relationship between pieces of information, each of which may need to meet multiple criteria. Deal with prominent competing information.</td><td>Integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. Compare, contrast or categorise taking many criteria into account. Deal with competing information.</td><td>Make connections or comparisons, give explanations, or evaluate a feature of text. Demonstrate a detailed understanding of the text in relation to familiar, everyday knowledge, or draw on less common knowledge.</td></tr>
<tr><td></td><td colspan="3">Continuous texts: Use conventions of text organisation, where present, and follow implicit or explicit logical links such as cause and effect relationships across sentences or paragraphs in order to locate, interpret or evaluate information.<br>Non-continuous texts: Consider one display in the light of a second, separate document or display, possibly in a different format, or combine several pieces of spatial, verbal and numeric information in a graph or map to draw conclusions about the information represented.</td></tr>
<tr><td>Level 2</td><td>Locate one or more pieces of information, each of which may be required to meet multiple criteria. Deal with competing information.</td><td>Identify the main idea in a text, understand relationships, form or apply simple categories, or construe meaning within a limited part of the text when the information is not prominent and low-level inferences are required.</td><td>Make a comparison or connections between the text and outside knowledge, or explain a feature of the text by drawing on personal experience and attitudes.</td></tr>
<tr><td></td><td colspan="3">Continuous texts: Follow logical and linguistic connections within a paragraph in order to locate or interpret information; or synthesise information across texts or parts of a text in order to infer the author's purpose.<br>Non-continuous texts: Demonstrate a grasp of the underlying structure of a visual display such as a simple tree diagram or table, or combine two pieces of information from a graph or table.</td></tr>
<tr><td>Level 1</td><td>Locate one or more independent pieces of explicitly stated information, typically meeting a single criterion, with little or no competing information in the text.</td><td>Recognise the main theme or author's purpose in a text about a familiar topic, when the required information in the text is prominent.</td><td>Make a simple connection between information in the text and common, everyday knowledge.</td></tr>
<tr><td></td><td colspan="3">Continuous texts: Use redundancy, paragraph headings or common print conventions to form an impression of the main idea of the text, or to locate information stated explicitly within a short section of text.<br>Non-continuous texts: Focus on discrete pieces of information, usually within a single display such as a simple map, a line graph or a bar graph that presents only a small amount of information in a straightforward way, and in which most of the verbal text is limited to a small number of words or phrases.</td></tr>
</table>





at each level build on and encompass the proficiencies laid down in the next lower level. This means that a student who is judged to be at Level 3 on a reading literacy scale is proficient not only in Level 3 tasks but also in Level 1 and 2 tasks. This also means that students who are at Levels 1 and 2 will be expected to get the average Level 3 item correct less than 50% of the time. Put another way, they will be expected to score less than 50% on a test made up of items drawn from Level 3.

Figure 2.10 shows the probability that individuals performing at selected points along the combined reading literacy scale will give a correct response to tasks of varying difficulty. One is a Level 1 task, one is a Level 3 task, and the third task receives two score points: one at Level 4 and the other at Level 5. It is readily seen here that a student with a score of 298, who is estimated to be below Level 1, has only a 43% chance of responding correctly to the Level 1 task that is at 367 on the reading literacy scale. This person has only a 14% chance of responding to the item from Level 3 and almost no chance of responding correctly to the item from Level 5. Someone with a proficiency of 371, in the middle of Level 1, has a 63% chance of responding to the item at 367, but only slightly more than one chance in four of responding correctly to the task at 508, and only a seven% chance of responding correctly to the task selected from Level 5. In contrast, someone at Level 3 would be expected to respond correctly 89% of the time to tasks at 367 on the reading literacy scale, and 64% of the time to tasks at 508, near the middle of Level 3. However, he or she would only have just over one chance in four (27%) of correctly responding to items from the middle of Level 5. Finally, a student at Level 5 is expected to respond correctly most of the time to almost all the tasks. As shown in Figure 2.10, a student having a score of 662 on the combined reading literacy scale has a 98% chance of answering the task at 367 correctly, a 90% chance of answering the item at Level 3 (508) correctly and a 65% of responding correctly to the task selected from near the centre of Level 5 (652).

Figure 2.10 also implicitly raises questions concerning the highest and lowest designated levels. Even though the top of the reading literacy scale is unbounded, it can be stated with some certainty that students of extremely high proficiency are capable of performing tasks characterised by the highest level of proficiency. There is more of an issue for students who are at the bottom end of the reading literacy scale. Level 1 begins at 335, yet a certain percentage of students in each country is estimated to be below this point on the scale. While there are no reading literacy tasks with a scale

Figure 2.10 ■ **Probability of responding correctly to selected tasks of varying difficulty for students with varying levels of proficiency**

| | **Selected tasks of varying difficulty:** | | | |
|---|---|---|---|---|
| **Students with varying levels of proficiency** | **Level 1 item at 367 points** | **Level 3 item at 508 points** | **Level 4 item at 567 points** | **Level 5 item at 652 points** |
| Below Level 1 (Proficiency of 298 points) | 43 | 14 | 8 | 3 |
| Level 1 (Proficiency of 371 points) | 63 | 27 | 16 | 7 |
| Level 2 (Proficiency of 444 points) | 79 | 45 | 30 | 14 |
| Level 3 (Proficiency of 517 points) | 89 | 64 | 48 | 27 |
| Level 4 (Proficiency of 589 points) | 95 | 80 | 68 | 45 |
| Level 5 (Proficiency of 662 points) | 98 | 90 | 82 | 65 |



value below 335, it is not correct to say that these students are without any reading literacy skills or are totally illiterate. However, on the basis of their performance in the set of tasks used in this assessment, they would be expected to score less than 50% on a set of tasks selected from Level 1. They are classified, therefore, as performing below Level 1.

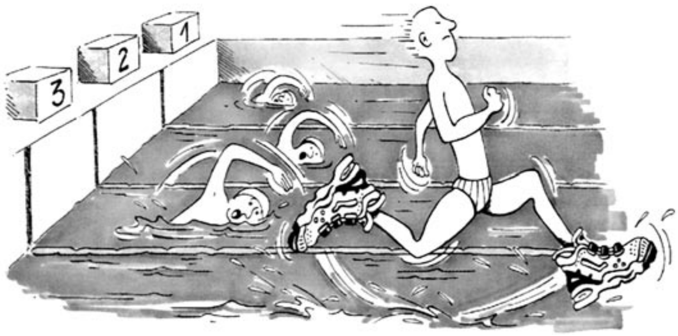
Since comparatively few young adults in our societies have no literacy skills, the framework does not call for a measure of whether or not 15-year-old students can read in a technical sense. That is, PISA does not measure the extent to which 15-year-old students are fluent readers or how competent they are at word recognition tasks or spelling. It does, however, reflect the contemporary view that students should, upon completing compulsory education, be able to construct, extend and reflect on the meaning of what they have read across a wide range of continuous and non-continuous texts commonly associated with a variety of situations both within and outside school. While it was not possible to say what knowledge and skills students performing below Level 1 may possess with regard to *reading literacy*, their level of proficiency indicates that these students are unlikely to be able to use reading independently as a tool to assist them in acquiring knowledge and skills in other areas.

## READING EXAMPLES

# Reading Example 1: RUNNERS

### Feel good in your runners

For 14 years the Sports Medicine Centre of Lyon (France) has been studying the injuries of young sports players and sports professionals. The study has established that the best course is prevention … and good shoes.



Knocks, falls, wear and tear...

Eighteen per cent of sports players aged 8 to 12 already have heel injuries. The cartilage of a footballer's ankle does not respond well to shocks, and 25% of professionals have discovered for themselves that it is an especially weak point. The cartilage of the delicate knee joint can also be irreparably damaged and if care is not taken right from childhood (10–12 years of age), this can cause premature osteoarthritis. The hip does not escape damage either and, particularly when tired, players run the risk of fractures as a result of falls or collisions.

According to the study, footballers who have been playing for more than ten years have bony outgrowths either on the tibia or on the heel. This is what is known as "footballer's foot", a deformity caused by shoes with soles and ankle parts that are too flexible.

Protect, support, stabilise, absorb

If a shoe is too rigid, it restricts movement. If it is too flexible, it increases the risk of injuries and sprains. A good sports shoe should meet four criteria:

Firstly, it must **provide exterior protection**: resisting knocks from the ball or another player, coping with unevenness in the ground, and keeping the foot warm and dry even when it is freezing cold and raining.

It must **support the foot**, and in particular the ankle joint, to avoid sprains, swelling and other problems, which may even affect the knee.

It must also provide players with good **stability** so that they do not slip on a wet ground or skid on a surface that is too dry.

Finally, it must **absorb shocks**, especially those suffered by volleyball and basketball players who are constantly jumping.

Dry feet

To avoid minor but painful conditions such as blisters or even splits or athlete's foot (fungal infections), the shoe must allow evaporation of perspiration and must prevent outside dampness from getting in. The ideal material for this is leather, which can be water-proofed to prevent the shoe from getting soaked the first time it rains.

**Source:** Revue ID (16) 1-15 June 1997.

RUNNERS is a piece of expository prose from a French-Belgian magazine produced for adolescent students. It is classed as belonging to the educational situation. One of the reasons for its selection as part of the reading instrument is its subject, which is considered of high interest for the PISA population of 15-year-old students. The article includes an attractive cartoon-like illustration and is broken up with catchy sub-headings. Within the continuous text format category, it is an example of expository writing in that it provides an outline of a mental construct, laying out a set of criteria for judging the quality of running shoes in terms of their fitness for young athletes.

The four tasks based on this stimulus covered all three aspects – retrieving information, interpreting and reflecting – but all are relatively easy, falling within Level 1. One of the tasks is reproduced below.



**Question 1: RUNNERS**

*According to the article, why should sports shoes not be too rigid?*

Score 1 (392).

*Answers which refer to restriction of movement.*

This task is classified as retrieving information in terms of aspect. It requires readers to take account of a single criterion to locate a piece of explicitly stated information.

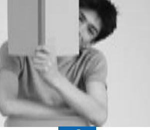
One factor that contributes to the difficulty of a task is the extent to which the wording of the question matches wording in the text. In this example, the reader can directly match on the word "rigid", which appears in the question and in the relevant part of the text, making the information easy to find.

Another factor that contributes to the difficulty of a task is the location and prominence of the information in the text: for example, information near the beginning of a text is typically easiest to find. Although the required information for this task is midway through the text, it is quite prominent because it is near the beginning of one of the three sections marked by sub-headings.

Another reason that this task is relatively easy is that full credit can be gained by quoting directly from the text: "it restricts movement". Many students, however, used their own words, such as "They prevent you from running easily" or "So you can move around".

A common error was to give an answer such as "Because you need support for your foot" – the opposite of the answer required, though it is also an idea located in the text. Students who gave this kind of answer may have overlooked the negative in the question ("… not be too rigid"), or made their own association between the ideas of "rigidity" and "support", leading them to a section of the text that was not relevant to this task. Other than this, there is little competing information to distract the reader.

---





## Reading Example 2: LAKE CHAD

Figure A shows changing levels of Lake Chad, in Saharan North Africa. Lake Chad disappeared completely in about 20 000 BC, during the last Ice Age. In about 11 000 BC it reappeared. Today, its level is about the same as it was in AD 1000.

Figure A

Lake Chad: changing levels

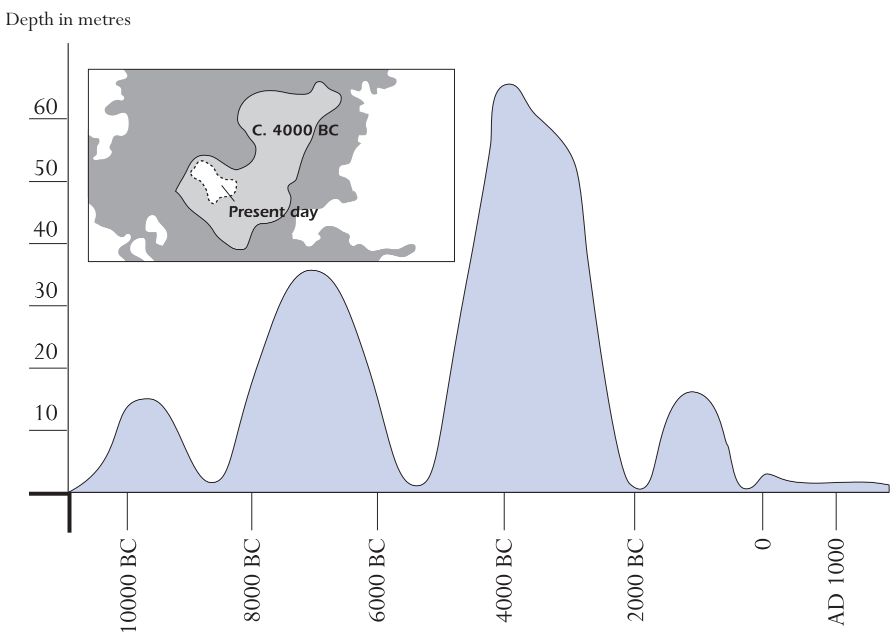
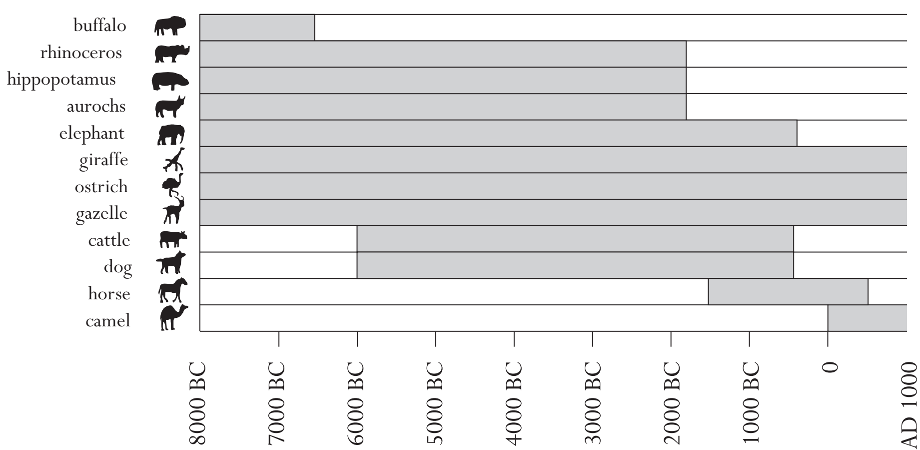



Figure B shows Saharan rock art (ancient drawings or paintings found on the walls of caves) and changing patterns of wildlife.

Figure B

Saharan rock art and changing patterns of wildlife




**Source:** Copyright Bartholomew Ltd. 1988. Extracted from **The Times Atlas of Archaeology** and reproduced by permission of Harper Collins Publishers.



The LAKE CHAD unit presents two graphs from an archaeological atlas. Figure A in LAKE CHAD is a line graph, and Figure B is a horizontal histogram. A third non-continuous text type is represented by a small map of the lake embedded in Figure A. Two very short passages of prose are also part of the stimulus, but as the tasks related to this stimulus are mostly related to the non-continuous components, they are classified as non-continuous on the text format dimension.

By juxtaposing several pieces of information the author invites the reader to infer a connection between the changing water levels of Lake Chad over time, and the periods in which certain species of wildlife inhabited its surroundings.

This is a type of text that might typically be encountered by students in an educational setting. Nevertheless, because the atlas is published for the general reader the text is classified as public in the situation dimension of the reading framework. The full set of five tasks that accompanied this stimulus covered all three aspects. One task, illustrating the Interpreting aspect, is reproduced below.

---

### Question 1: LAKE CHAD

For this question you need to draw together information from Figure 1 and Figure 2.

The disappearance of the rhinoceros, hippopotamus and aurochs from Saharan rock art happened

A at the beginning of the most recent Ice Age.

B in the middle of the period when Lake Chad was at its highest level.

C after the level of Lake Chad had been falling for over a thousand years.

D at the beginning of an uninterrupted dry period.

Score 1 (508)

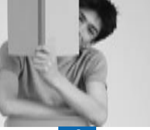
The correct answer is option C.

This interpreting task requires students to integrate several parts of the non-continuous texts to understand a relationship. They need to compare information given in two graphs.

The requirement to combine information from two sources contributes to the task's moderate difficulty. An added contributor to its difficulty is that two different types of graphs are used (a line graph and a histogram), and the reader needs to have interpreted the structure of both in order to translate the relevant information from one form to the other.

Of those students who did not select the correct answer, the largest proportion chose distractor D, "at the beginning of an uninterrupted dry period." If one disregards the texts, this seems the most plausible of the wrong answers, and its popularity indicates that the students who chose it might have been drawing on familiar knowledge from outside the text, rather than on the information in front of them.

---





## Reading Example 3: GRAFFITI

I'm simmering with anger as the school wall is cleaned and repainted for the fourth time to get rid of graffiti. Creativity is admirable but people should find ways to express themselves that do not inflict extra costs upon society.

Why do you spoil the reputation of young people by painting graffiti where it's forbidden? Professional artists do not hang their paintings in the streets, do they? Instead they seek funding and gain fame through legal exhibitions.

In my opinion buildings, fences and park benches are works of art in themselves. It's really pathetic to spoil this architecture with graffiti and what's more, the method destroys the ozone layer. Really, I can't understand why these criminal artists bother as their "artistic works" are just removed from sight over and over again.

**Helga**

There is no accounting for taste. Society is full of communication and advertising. Company logos, shop names. Large intrusive posters on the streets. Are they acceptable? Yes, mostly. Is graffiti acceptable? Some people say yes, some no.

Who pays the price for graffiti? Who is ultimately paying the price for advertisements? Correct. The consumer.

Have the people who put up billboards asked your permission? No. Should graffiti painters do so then? Isn't it all just a question of communication – your own name, the names of gangs and large works of art in the street?

Think about the striped and chequered clothes that appeared in the stores a few years ago. And ski wear. The patterns and colours were stolen directly from the flowery concrete walls. It's quite amusing that these patterns and colours are accepted and admired but that graffiti in the same style is considered dreadful.

Times are hard for art.

**Sophia**

**Source:** Mari Hamkala.

The stimulus for this unit, originally from Finland, consists of two letters posted on the Internet. The four tasks that accompanied the stimulus simulate typical literacy activities, since as readers we often synthesise, compare and contrast ideas from two or more different sources.

Because they were published on the Internet, the GRAFFITI letters are classified as public in terms of situation. They are classified as argumentation within the broader classification of continuous texts, as they set forth propositions and attempt to persuade the reader to a point of view.

As with RUNNERS, the subject matter of GRAFFITI was expected to be interesting for 15-year-old students: the implied debate between the writers as to whether graffiti makers are artists or vandals would represent a real issue in the minds of the test-takers.

One of the tasks, representing the aspect reflection and evaluation**,** is reproduced below.

### Question 1: GRAFFITI

We can talk about *what* a letter says (its content).

We can talk about *the way* a letter is written (its style).

Regardless of which letter you agree with, in your opinion, which do you think is the better letter? Explain your answer by referring to *the way* one or both letters are written.

Score 1 (581)

Answers which explain opinion with reference to the style or form of one or both letters. They should refer to criteria such as style of writing, structure of argument, cogency of argument, tone, register, or strategies for persuading readers. Terms like "better argu-ments" must be substantiated.

*This task requires students to use formal knowledge to evaluate the writer's craft by comparing the two letters. In the five-aspect categorisation, this task is classified as reflecting on the form of a text, since to answer it readers need to draw on their own understanding of what constitutes good writing.*

*Full credit was given to many types of answers, including those dealing with one or both writers' tone or argumentative strategies, or with the structure of the piece. Some typical answers that earned full credit were: "Helga's letter was effective because of the way she addressed the graffiti artists directly" or "In my opinion, the second letter is better because it has questions that involve you making you feel that you are having a discussion rather than a lecture".*

*Answers that were not given credit were often vague, or offered a general opinion without substantiating it with reference to the text, or related to content rather than style (for example, "Sophia, because graffiti is a form of art.")*

---

## SUMMARY

The notion of reading literacy in PISA goes beyond the simple measurement of a student's capacity to decode and comprehend literal information. Reading literacy in PISA involves also understanding, using and reflecting on written texts. It also takes into account the importance of reading literacy in achieving goals and participating as an active citizen in society.

There is a recognition that students engage in reading in different ways. PISA makes a distinction between continuous text, such articles that a student might read in a magazine, a newspaper or a novel and non-continuous text such as charts, tables, maps and diagrams. Students are also presented with a variety of item types including multiple choice, open- and closed- constructed responses.

Reading literacy in PISA is reported on three subscales – retrieving information, interpreting text, and reflecting on and evaluating text. Following the PISA 2000 assessment, five proficiency levels were developed to indicate students' capacities in the reading assessment. Students at the highest level are able to carry out high order tasks such as locating complex information from within an unfamiliar text which contains competing information, whereas at the lowest proficiency levels students are only able to locate information which is more evident and has less competing information accompanying it. At the highest levels students are expected to reflect on the goals of an author in a certain piece of text, whereas students at the lower levels would be expected to make a simple connection between information in the text and everyday life.

Reading was the main domain of the first cycle of PISA and will be again in 2009, at which time the assessment framework will undergo a review to consider developments that have occurred in that time.

3

# Mathematical Literacy



## DEFINITION OF THE DOMAIN

The PISA *mathematical literacy* domain is concerned with the capacities of students to analyse, reason and communicate ideas effectively as they pose, formulate, solve and interpret mathematical problems in a variety of situations. The PISA assessment focuses on real-world problems, moving beyond the kinds of situations and problems typically encountered in school classrooms. In real-world settings, citizens regularly face situations when shopping, travelling, cooking, dealing with their personal finances, judging political issues, etc., in which the use of quantitative or spatial reasoning or other mathematical competencies would help clarify, formulate or solve a problem. Such uses of mathematics are based on the skills learned and practised through the kinds of problems that typically appear in school textbooks and classrooms. However, they demand the ability to apply those skills in a less structured context, where the directions are not so clear, and where the student must make decisions about what knowledge may be relevant and how it might usefully be applied.

PISA *mathematical literacy* deals with the extent to which 15-year-old students can be regarded as informed, reflective citizens and intelligent consumers. Citizens in every country are increasingly confronted with a myriad of tasks involving quantitative, spatial, probabilistic or other mathematical concepts. For example, media outlets (newspapers, magazines, television and the Internet) are filled with information in the form of tables, charts and graphs about such subjects as weather, economics, medicine and sports, to name a few. Citizens are bombarded with information on issues such as global warming and the greenhouse effect, population growth, oil slicks and the seas or the disappearing countryside. Last but not least, citizens are confronted with the need to read forms, to interpret bus and train timetables, to successfully carry out transactions involving money, to determine the best buy at the market, etc. PISA *mathematical literacy* focuses on the capacity of 15-year-old students (the age when many students are completing their formal compulsory mathematics learning) to use their mathematical knowledge and understanding to help make sense of these issues and to carry out the resulting tasks.

PISA defines *mathematical literacy* as:

> *Mathematical literacy* is an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen.

Some explanatory remarks may help to further clarify this domain definition:

- The term "mathematical literacy" emphasises mathematical knowledge put to functional use in a multitude of different situations in varied, reflective and insight-based ways. Of course, for such use to be possible and viable, many fundamental mathematical knowledge and skills are needed. Literacy in the linguistic sense presupposes, but cannot be reduced to, a rich vocabulary and a substantial knowledge of grammatical rules, phonetics, orthography, etc. To communicate, humans combine these elements in creative ways in response to each real-world situation encountered. In the same way, *mathematical literacy* cannot be reduced to, but certainly presupposes, knowledge of mathematical terminology, facts and procedures, as well as skills in performing certain operations and carrying out certain methods. *Mathematical literacy* involves the creative combining of these elements in response to the demands imposed by the external situation.



- The term "the world" means the natural, social and cultural setting in which the individual lives. As Freudenthal (1983) stated: "Our mathematical concepts, structures, ideas have been invented as tools to organise the phenomena of the physical, social and mental world" (p. ix).
- The term "to use and engage with" is meant to cover using mathematics and solving mathematical problems, and also implies a broader personal involvement through communicating, relating to, assessing, and even appreciating and enjoying mathematics. Thus the definition of *mathematical literacy* encompasses the functional use of mathematics in a narrow sense as well as preparedness for further study and the aesthetic and recreational elements of mathematics.
- The phrase "that individual's life" includes his or her private life, occupational life and social life with peers and relatives, as well as his or her life as a citizen of a community.

A crucial capacity implied by this notion of *mathematical literacy* is the ability to pose, formulate, solve and interpret problems using mathematics within a variety of situations or contexts. The contexts range from purely mathematical ones to contexts in which no mathematical structure is present or apparent at the outset – the problem poser or solver must successfully introduce the mathematical structure. It is also important to emphasise that the definition is not just concerned with knowing mathematics at some minimal level; it is also about doing and using mathematics in situations that range from the everyday to the unusual, from the simple to the complex.

Mathematics-related attitudes and emotions such as self-confidence, curiosity, feelings of interest and relevance, and the desire to do or understand things, are not components of the definition of *mathematical literacy* but nevertheless are important contributors to it. In principle it is possible to possess *mathematical literacy* without possessing such attitudes and emotions. In practice, however, it is not likely that such literacy is going to be exerted and put into practice by someone who does not have some degree of self-confidence, curiosity, feelings of interest and relevance, and the desire to do or understand things that contain mathematical components. The importance of these attitudes and emotions as correlates of *mathematical literacy* is recognised. They are not part of the mathematical literacy assessment, but will be addressed in other components of PISA.

## THEORETICAL BASIS FOR THE PISA MATHEMATICS FRAMEWORK

The PISA definition of *mathematical literacy* is consistent with the broad and integrative theory about the structure and use of language as reflected in recent socio-cultural literacy studies. In James Gee's *Preamble to a Literacy Program* (1998), the term "literacy" refers to the human use of language. The ability to read, write, listen and speak a language is the most important tool through which human social activity is mediated. In fact, each human language and use of language has an intricate design tied in complex ways to a variety of functions. For a person to be literate in a language implies that the person knows many of the design resources of the language and is able to use those resources for several different social functions. Analogously, considering mathematics as a language implies that students must learn the design features involved in mathematical discourse (the terms, facts, signs and symbols, procedures and skills in performing certain operations in specific mathematical subdomains, and the structure of those ideas in each sub-domain), and they also must learn to use such ideas to solve non-routine problems in a variety of situations defined in terms of social functions. Note that the design features for mathematics include knowing the basic terms, procedures and concepts commonly taught in schools, and also involve knowing how these features are structured and used. Unfortunately, one can know a good deal about the design



features of mathematics without knowing either their structure or how to use those features to solve problems. These scholarly notions involving the interplay of design features and functions that support the mathematics framework for PISA can be illustrated via the following example.

## Mathematics Example 1: HEARTBEAT

For health reasons people should limit their efforts, for instance during sports, in order not to exceed a certain heartbeat frequency.

For years the relationship between a person's recommended maximum heart rate and the person's age was described by the following formula:

Recommended maximum heart rate = 220 – age

Recent research showed that this formula should be modified slightly. The new formula is as follows:

Recommended maximum heart rate = 208 – (0.7 × age)

The questions in this unit centre around the difference between the two formulas and how they affect the calculation of the maximum allowable heart rate.

This problem can be solved by following the general strategy used by mathematicians, which the mathematics framework will refer to as "mathematising". Mathematising can be characterised as having five aspects:

- *In the first step, the process of mathematisation or mathematising starts with a problem situated in reality*

  As will be clear from the item the reality in this case is physical health and fitness: "An important rule when exercising is that one should be careful not push oneself too far as excessive exertion may cause health problems." The question alerts us to this issue through the text linking health to heart rate and by referring to "recommended maximum heart rate".

- *In the second step, the problem solver tries to identify the relevant mathematics, and reorganises the problem according to the mathematical concepts identified*

  It seems clear that the student faces two word formulas that need to be understood and he or she is requested to compare these two formulas, and what they really mean in mathematical terms. The formulas give a relation between the advised maximum heart beat rate and the age of a person.

- *The third step involves gradually trimming away the reality*

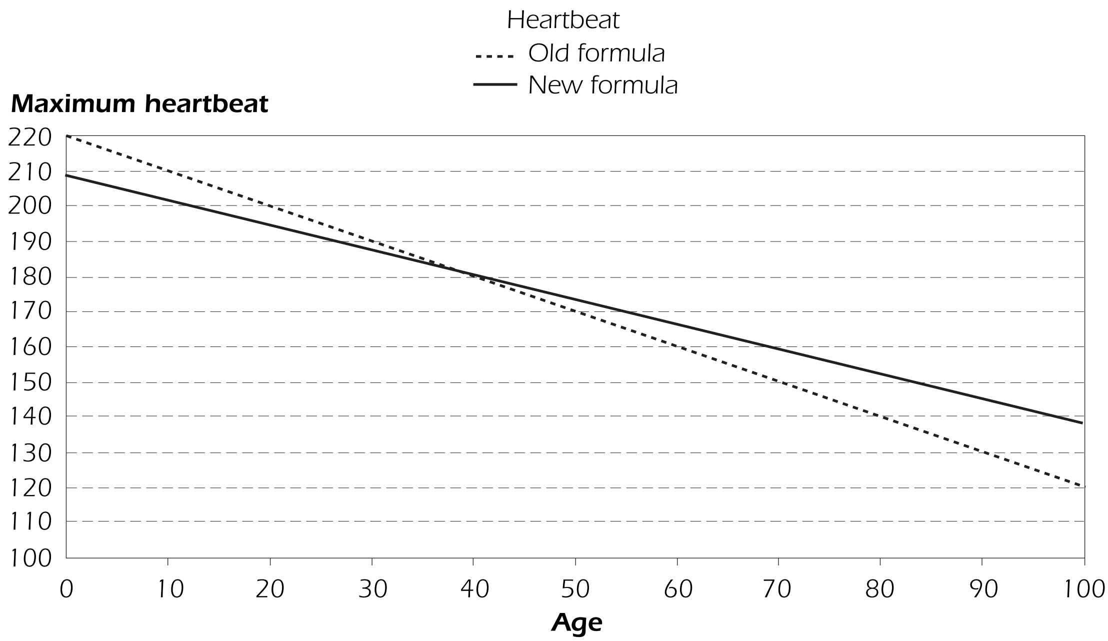
  There are different ways to move the problem to a strictly mathematical problem, or of trimming away reality. One way to go is to make the word formulas into more formal algebraic expressions like $y = 220 - x$ or $y = 208 - 0.7\,x$. The student must remember that $y$ expresses the maximum heart beat in beats per minute and $x$ represents the age in years. Another strictly mathematical approach would be to draw the graphs directly from the word formulas. These graphs are straight lines as the formulas are of the first degree. The graphs have different slopes, so they intersect:

  These three steps lead us from a real world problem to a mathematical problem.

- *The fourth step is solving the mathematical problem*

  The mathematical problem at hand is to compare the two formulas or graphs, and to say something about the differences for people of a certain age. A nice way to start is to find out where the two formulas give equal results or where the two graphs intersect. So the student can find this by solving the equation: $220 - x = 208 - 0.7x$. This gives us $x = 40$ and the corresponding value for $y$ is 180. So the two graphs intersect at the point (40, 180).

This point can also be found in the graph below. As the slope of the first formula is -1 and the second is -0.7 the student knows that the second graph is less steep than the first one. Or the graph of $y = 220 - x$ lies above the graph of $y = 208 - 0.7x$ for values of $x$ smaller than 40 and lies below for values of $x$ larger than 40.




- *The fifth step asks what the meaning of the strictly mathematical solution is in terms of the real world*

The meaning is not too difficult if the student realises that $x$ is the age of a person and y the maximum heart beat. If one is 40 years old both formulas give the same result: a maximum heartbeat of 180. The 'old' rule allows for higher heart rates for younger people: in the extreme, if the age is zero the maximum is 220 in the old formula and only 208 in the new formula. But for older people, in this case for those over 40, the more recent insights allow for higher maximum heartbeat; as an example: for an age of 100 years the student sees that the old formula gives him or her a maximum of 120 and the new one 138. Of course the student has to realise a number of other things: the formulas lack mathematical precision and give a feel of being only quasi scientific. In reality, the formulas provide only a rule of thumb that should be used with caution. Another point is that for ages at the extreme the outcomes should be taken with even more caution.

What this example shows is that even with items that are relatively simple in the sense that they can be used within the restrictions of a large international study and can be solved in a short time, the full cycle of mathematisation and problem solving can still be identified.

These processes characterise how, in a broad sense, mathematicians often do mathematics, how people use mathematics in a variety of current and potential occupations, and how informed and reflective citizens should use mathematics to fully and competently engage with the real world. In fact, learning to mathematise should be a primary educational goal for all students.



Today and in the foreseeable future, every country needs mathematically literate citizens to deal with a very complex and rapidly changing society. Accessible information has been growing exponentially and citizens need to be able to decide how to deal with this information. Social debates increasingly involve quantitative information to support claims. One example of the need for *mathematical literacy* is the frequent demand for individuals to make judgements and assess the accuracy of conclusions and claims in surveys and studies. Being able to judge the soundness of the claims from such arguments is, and increasingly will be, a critical aspect of being a responsible citizen. The steps of the mathematisation process discussed in this framework are the fundamental elements of using mathematics in such complex situations. Failure to use mathematical notions can result in confused personal decisions, an increased susceptibility to pseudo-sciences, and poorly informed decision-making in professional and public life.

A mathematically literate citizen realises how quickly change is taking place and the consequent need to be open to lifelong learning. Adapting to these changes in a creative, flexible and practical way is a necessary condition for successful citizenship. The skills learned at school will probably not be sufficient to serve the needs of citizens for the majority of their adult life.

The requirements for competent and reflective citizenship also affect the workforce. Workers are less and less expected to carry out repetitive physical chores. Instead, they are engaged actively in monitoring output from a variety of high-technology machines, dealing with a flood of information and engaging in team problem solving. The trend is that more and more occupations will require the ability to understand, communicate, use and explain concepts and procedures based on mathematical thinking. The steps of the mathematisation process are the building blocks of this kind of mathematical thinking.

Finally, mathematically literate citizens also develop an appreciation for mathematics as a dynamic, changing and relevant discipline that may often serve their needs.

The operational problem faced by PISA is how to assess whether 15-year-old students are mathematically literate in terms of their ability to mathematise. Unfortunately, in a timed assessment this is difficult because for most complex real situations the full process of proceeding from reality to mathematics and back often involves collaboration and finding appropriate resources, and takes considerable time.

To illustrate mathematisation in an extended problem-solving-exercise, consider the following *HOLIDAY* example, which was an item in the PISA 2003 problem-solving survey. The problem poses two questions to the students. It deals with the planning of a route and places to stay overnight on a holiday trip. Students were presented a simplified map and a chart (multiple representations) showing the distances between the towns illustrated on the map.



## Mathematics Example 2: HOLIDAY

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

Figures A and B show a map of the area and the distances between towns.

**Figure A. Map of roads between towns**

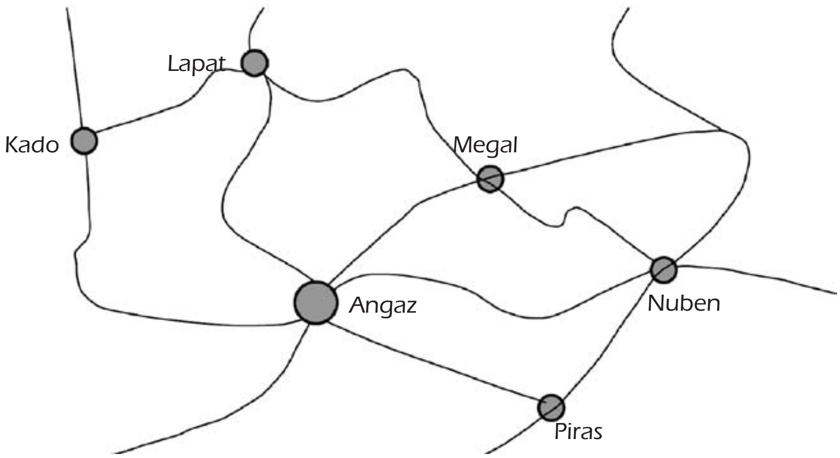



**Figure B. Shortest road distance of towns from each other in kilometres**

| | Angaz | Kado | Lapat | Megal | Nuben | Piras |
|---|---|---|---|---|---|---|
| Angaz | | | | | | |
| Kado | 550 | | | | | |
| Lapat | 500 | 300 | | | | |
| Mergal | 300 | 850 | 550 | | | |
| Nuben | 500 | | 1300 | 450 | | |
| Piras | 300 | 850 | 800 | 600 | 250 | |

### Question 1: HOLIDAY

Calculate the shortest distance by road between Nuben and Kado.

Distance: ......................................... kilometres.



**Question 2: HOLIDAY**

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

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

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

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

It can be noted that there is no obvious link to a curricular discipline, although there is a clear relation to discrete mathematics. There is also not a pre-described strategy to solve the problem. Often if problems are posed to students they know exactly which strategy to use. But in real world problem solving there is no well-known strategy available.

Furthermore, the five aspects of mathematising are clearly visible: the problem is situated in reality; it can be organised according to mathematical concepts (distance tables or matrices) and maps (as models of reality). The student needs to trim away the redundant information and focus on the relevant information, especially the mathematical aspects of that information. After solving the problem in mathematical terms the student needs to reflect on the solution in terms of the real situation.

Although there is relatively little reading required to solve the problem it still is rather complex because of the fact that students must read and interpret information from the map and from the distance chart.  Some of the distances that they have to find in the chart require them to read distances starting from the bottom of the chart, rather than from the left down. For example, in determining the distance from Nuben to Piras, one needs to transform the search to that of finding the distance from Piras to Nuben (*Problem Solving for Tomorrow's World – First Measures of Cross-Curricular Competencies from PISA 2003* [OECD, 2004]).

The second question sets a number of constraints that needed to be complied with simultaneously – a maximum of 300 km travelled in a given day, starting and finishing in Zoe's hometown of Angaz, visiting Kado and Lapat, and staying two nights in each of these cities so that she can achieve her vacation goals.

It should be noted than in the PISA problem-solving survey from which this item was taken, students were allowed considerably more time to find answers than the average time allowed for the mathematics items, which are typically shorter.

Ideally, to judge whether 15-year-old students can use their accumulated mathematical knowledge to solve mathematical problems they encounter in their world, one would collect information about



their ability to mathematise such complex situations. Clearly this is impractical. Instead, PISA has chosen to prepare items to assess different parts of this process. The following section describes the strategy chosen to create a set of test items in a balanced manner so that a selection of these items cover the five aspects of mathematising. The aim is to use the responses to those items to locate students on a scale of proficiency in the PISA construct of *mathematical literacy*.

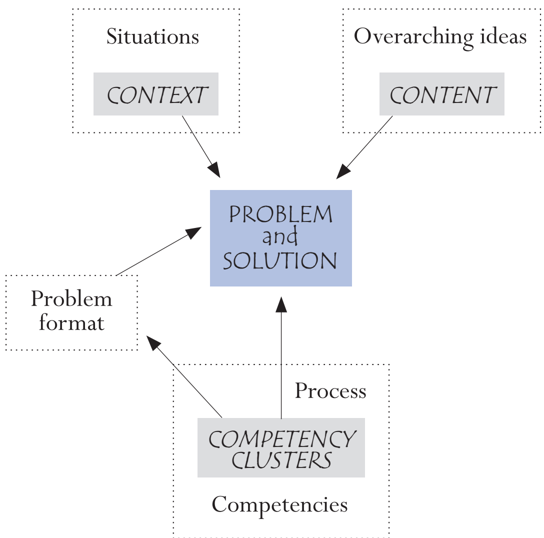
## ORGANISATION OF THE DOMAIN

The PISA mathematics framework provides the rationale for, and the description of, an assessment of the extent to which 15-year-old students can handle mathematics in a well-founded manner when confronted with real-world problems, or, in more general terms, an assessment of how mathematically literate 15-year-old students are. To describe the domain that is assessed more clearly, three components must be distinguished:

- The *situations or contexts* in which the problems are located.
- The *mathematical content* that has to be used to solve the problems, organised by certain overarching ideas
- The *competencies* that have to be activated in order to connect the real world, in which the problems are generated, with mathematics, and thus to solve the problems.

These components are represented in Figure 3.1. An explanation of each is provided afterwards.

Figure 3.1 ■ **The components of the mathematics domain**




The extent of a person's *mathematical literacy* is seen in the way he or she uses mathematical knowledge and skills in solving problems. Problems (and their solutions) may occur in a variety of situations or contexts within the experience of an individual. PISA problems draw from the real world in two ways. First, problems exist within some broad situations that are relevant to the student's life. The situations form part of the real world and are indicated by a big square in the upper left of the picture. Next, within that situation, problems have a more specific context. This is represented by the grey rectangle within the situations square.



In the *HEARTBEAT* and *HOLIDAY* examples above, the situation is the personal real world, and the contexts are sport/health aspects for the active citizen and how to plan a holiday.

The next component of the real world that has to be considered when thinking about *mathematical literacy* is the mathematical content that a person might bring to bear in solving a problem. The mathematical content can be illustrated by four categories that encompass the kinds of problems that arise through interaction with day-to-day phenomena, and that are based on a conception of the ways in which mathematical content presents itself to people. For PISA assessment purposes, these overarching ideas are: *space and shape*, *change and relationships*, *quantity* and *uncertainty*. This is somewhat different from an approach to content that would be familiar from the perspective of mathematics instruction and the curricular strands typically taught in schools. However, the overarching ideas together broadly encompass the range of mathematical topics that students are expected to have learned. The overarching ideas are represented by the big square in the upper right of the diagram in Figure 3.1. From the overarching ideas the content used in solving a problem is extracted. This is represented by the smaller square within the overarching ideas square.

The arrows going from the context and content to the problem show how the real world (including mathematics) makes up a problem.

The *HEARTBEAT* problem involves mathematical relations and comparing two relations in order to make decisions. So the problem belongs to the overarching idea *change and relationships*. The *HOLIDAY* problem requires some basic computation but the second question needs some analytic reasoning. The most appropriate overarching idea is *quantity*.

The mathematical processes that students apply as they attempt to solve problems are referred to as mathematical competencies. Three competency clusters encapsulate the different cognitive processes that are needed to solve various kinds of problems. These clusters reflect the way that mathematical processes are typically employed when solving problems that arise as students interact with their world, and will be described in detail in later sections.

Thus the process component of this framework is represented first by the large square, representing the general mathematical competencies, and a smaller square that represents the three competency clusters. The particular competencies needed to solve a problem are related to the nature of the problem, and the competencies used will be reflected in the solution found. This interaction is represented by the arrow from the competency clusters to the problem and its solution.

The remaining arrow goes from the competency clusters to the problem format. The competencies employed in solving a problem are related to the form of the problem and its precise demands.

It should be emphasised that the three components just described are of different natures. Indeed, the competencies are the core of *mathematical literacy*. Only when certain competencies are available to students will they be in a position to successfully solve given problems. Assessing *mathematical literacy* includes assessing to what extent students possess mathematical competencies they can productively apply in problem situations.

In the following sections, these three components are described in more detail.



## SITUATIONS AND CONTEXT

An important aspect of *mathematical literacy* is engagement with mathematics: using and doing mathematics in a variety of situations. It has been recognised that in dealing with issues that lend themselves to a mathematical treatment, the choice of mathematical methods and representations is often dependent on the situations in which the problems are presented.

The situation is the part of the student's world in which the tasks are placed. It is located at a certain distance from the students. For PISA, the closest situation is the student's personal life; next is school life, work life and leisure, followed by the local community and society as encountered in daily life. Furthest away are scientific situations. Four situation types are defined and used for problems to be solved: *personal, educational/occupational, public* and *scientific*.

The context of an item is its specific setting within a situation. It includes all the detailed elements used to formulate the problem.

Consider the following example:

## Mathematics Example 3: SAVINGS ACCOUNT

1 000 zed is put into a savings account at a bank. There are two choices: one can get an annual rate of 4% OR one can get an immediate 10 zed bonus from the bank, and a 3% annual rate.

---

### Question 1: SAVINGS ACCOUNT

Which option is better after one year? After two years?

The situation of this item is finance and banking, which is a situation from the local community and society that PISA would classify as public. The context of this item concerns money (zeds) and interest rates for a bank account.

Note that this kind of problem is one that could be part of the actual experience or practice of the participant in some real-world setting. It provides an authentic context for the use of mathematics, since the application of mathematics in this context would be genuinely directed to solving the problem. This can be contrasted with problems frequently seen in school mathematics texts, where the main purpose is to practise the mathematics involved rather than to use mathematics to solve a real problem. This authenticity in the use of mathematics is an important aspect of the design and analysis of items for PISA, strongly related to the definition of *mathematical literacy*.

Note that this use of the term "authentic" is not intended to indicate that mathematics items are necessarily in some sense genuine and real. PISA mathematics uses the term "authentic" to indicate that the use of mathematics is genuinely directed to solving the problem at hand, rather than the problem being merely a vehicle for the purpose of practising some mathematics.

It should also be noted that there are some made-up elements of the problem – the money involved is fictitious. This fictitious element is introduced to ensure that students from certain countries are not given an unfair advantage.

The situation and context of a problem can also be considered in terms of the distance between the problem and the mathematics involved. If a task refers only to mathematical objects, symbols



or structures, and makes no reference to matters outside the mathematical world, the context of the task is considered as intra-mathematical, and the task will be classified as belonging to the scientific situation type. A limited range of such tasks is included in PISA, where the close link between the problem and the underlying mathematics is made explicit in the problem context. More typically, problems encountered in the day-to-day experience of the student are not stated in explicit mathematical terms. They refer to real-world objects. These task contexts are extra-mathematical and the student must translate these problem contexts into a mathematical form. Generally speaking, PISA puts an emphasis on tasks that might be encountered in some real-world situation and possess an authentic context for the use of mathematics that influences the solution and its interpretation. Note that this does not preclude the inclusion of tasks in which the context is hypothetical, as long as the context has some real elements, is not too far removed from a real-world situation, and for which the use of mathematics to solve the problem would be authentic. Example 4 shows a problem with a hypothetical context that is extra-mathematical.

## Mathematics Example 4: COINAGE SYSTEM

### Question 1: COINAGE SYSTEM

Would it be possible to establish a coinage system based on only the denominations 3 and 5? More specifically, what amounts could be reached on that basis? Would such a system be desirable?

This problem derives its quality not primarily from its closeness to the real world, but from the fact that it is mathematically interesting and calls on competencies that are related to *mathematical literacy*. The use of mathematics to explain hypothetical scenarios and explore potential systems or situations, even if these are unlikely to be carried out in reality, is one of its most powerful features. Such a problem would be classified as belonging to the scientific situation type.

In summary, PISA places most value on tasks that could be encountered in a variety of real-world situations and that have a context in which the use of mathematics to solve the problem would be authentic. Problems with extra-mathematical contexts that influence the solution and its interpretation are preferred as a vehicle for assessing *mathematical literacy* since these problems are most like those encountered in day-to-day life.

### MATHEMATICAL CONTENT – THE FOUR OVERARCHING IDEAS

Today, many see mathematics as the science of patterns in a general sense. Thus, this framework has chosen overarching ideas that reflects this: patterns in *space and shape*, patterns in *change and relationships,* patterns in *quantity* form central and essential concepts for any description of mathematics, and they form the heart of any curriculum, at any level. But to be literate in mathematics means more. Dealing with uncertainty from a mathematical and scientific perspective is essential. For this reason, elements of probability theory and statistics give rise to the fourth overarching idea: *uncertainty*.

The following list of overarching ideas, therefore, is used in PISA 2006 to meet the requirements of historical development, coverage of the domain and reflection of the major threads of school curriculum:

- *Space and shape*
- *Change and relationships*
- *Quantity*
- *Uncertainty*



With these four, mathematical content is organised into a sufficient number of areas to ensure a spread of items across the curriculum, but at the same time a number small enough to avoid a too fine division that would work against a focus on problems based in real situations.

The basic conception of an overarching idea is an encompassing set of phenomena and concepts that make sense and can be encountered within and across a multitude of different situations. By its very nature, each overarching idea can be perceived as a sort of general notion dealing with some generalised content dimension. This implies that the overarching ideas cannot be sharply delineated vis-à-vis one another, nor can traditional mathematics content strands. Rather, each of them represents a certain perspective, or point of view, which can be thought of as possessing a core, a centre of gravity and somewhat blurred outskirts that allow for intersection with other overarching ideas. In principle, any overarching idea intersects any other overarching idea. The four overarching ideas are described in the following section.

### Space and shape

Patterns are encountered everywhere: in spoken words, music, video, traffic, building constructions and art. Shapes can be regarded as patterns: houses, office buildings, bridges, starfish, snowflakes, town plans, cloverleaves, crystals and shadows. Geometric patterns can serve as relatively simple models of many kinds of phenomena and their study is possible and desirable at all levels (Grünbaum, 1985).

It is also important to have an understanding of the properties of objects and their relative positions. Students must be aware of how and why they see things and must learn to navigate through space and through constructions and shapes. This means understanding the relationship between shapes and images or visual representations, such as that between a real city and photographs and maps of the same city. It includes also understanding how three-dimensional objects can be represented in two dimensions, how shadows are formed and must be interpreted, what perspective is and how it functions.

Shape has strong ties to traditional geometry but goes far beyond it in content, meaning and method. Interaction with real shapes involves understanding the visual world, its description and encoding and decoding of visual information. It also means interpretation of visual information. In order to grasp the concept of shapes, students should be able to discover the way in which objects are similar and how they differ, to analyse the different components of the object, and to recognise shapes in different dimensions and representations.

It is important to note that shapes can be more than static entities. A shape can be transformed as an entity and shapes can be modified. These changes can sometimes be visualised very elegantly using computer technology. Students should be able to see the patterns and regularities when shapes are changing. An example is shown in Figure 3.2 in the following section.

The study of shape and constructions requires looking for similarities and differences when analysing the components of form and recognising shapes in different representations and different dimensions. The study of shapes is closely connected to the concept of grasping space (Freudenthal, 1973).

Examples requiring this kind of thinking are abundant. Identifying and relating a photograph of a city to a map of that city and indicating from which point a picture was taken; the ability to draw a



map; understanding why a building nearby looks bigger than a building further away; understanding how the rails of a railway track appear to meet at the horizon – all these are relevant for students within this overarching idea.

As students live in a three-dimensional space, they should be familiar with views of objects from three orthogonal aspects (for example the front, the side and from above). They should be aware of the power and limitations of different representations of three-dimensional shapes as indicated by the example provided in the following Figure 3.3. Not only must they understand the relative position of objects, but also how they can navigate through space and through constructions and shapes. An example is reading and interpreting a map and designing instructions on how to get from point A to point B using coordinates, common language or a picture.

Conceptual understanding of shapes also includes the ability to take a three-dimensional object and make a two-dimensional net of it, and vice-versa, even if the three-dimensional object is presented in two dimensions. An example of this is given in Figure 3.4.

Key aspects of *space and shape* are:

- Recognising shapes and patterns
- Describing, encoding and decoding visual information
- Understanding dynamic changes to shapes
- Similarities and differences
- Relative positions
- 2-D and 3-D representations and the relations between them
- Navigation through space

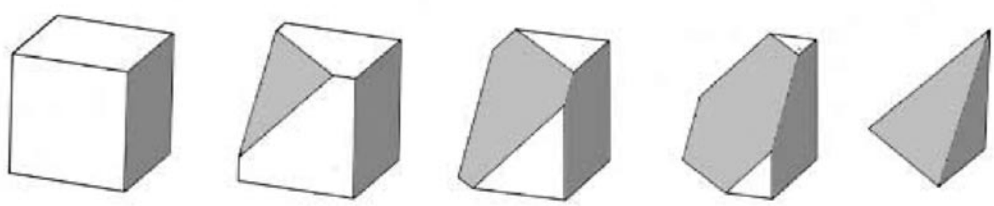
### Space and shape examples

Figure 3.2 shows a simple example of the need for flexibility in seeing shapes as they change. It is based on a cube that is being 'sectioned' (that is, plane cuts are made through the cube). A variety of questions could be asked, such as:

Figure 3.2 ■ **A cube, with plane cuts in various places**

What shapes can be produced by one plane cut through a cube?

How many faces, edges, or vertices will be produced when a cube is sectioned in this way?







Three examples of the need for familiarity with representations of three-dimensional shapes follow. In the first example, the side and front view of an object constructed of cubes is given in Figure 3.3. The question is:

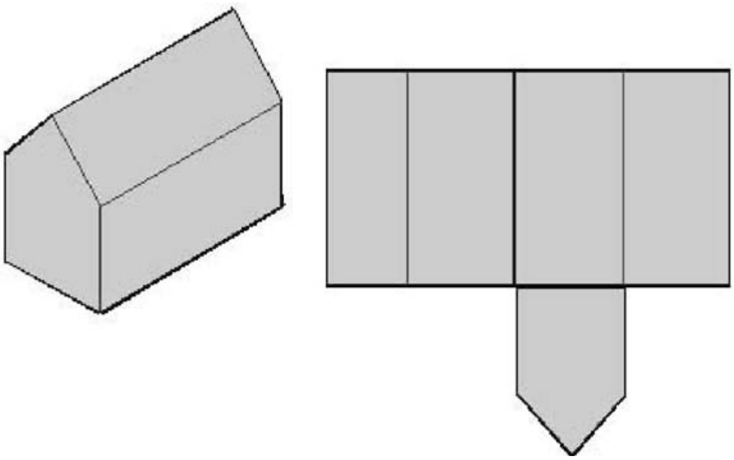
Figure 3.3 ■ **Side and front views of an object made from cubes**

How many cubes have been used to make this object?


It may come as a surprise to many – students and teachers alike – that the maximum number of cubes is 20 and the minimum is 6 (de Lange, 1995).

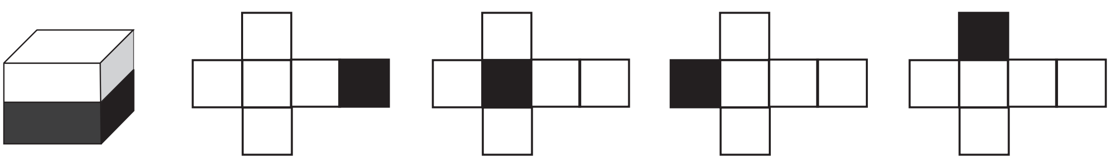
The next example shows a two-dimensional representation of a barn and an incomplete net of the barn. The problem is to complete the net of the barn.

Figure 3.4 ■ **Two-dimensional representation of a three-dimensional barn and its (incomplete) net**



A final example similar to the previous one is shown in Figure 3.5 (adapted from Hershkovitz *et al.*, 1996).

Figure 3.5 ■ **Cube with black bottom**





The lower half of the cube has been painted black. For each of the four nets, the bottom side is already black. Students could be asked to finish each net by shading the right squares.

### Change and relationships

Every natural phenomenon is a manifestation of change and the world around us displays a multitude of temporary and permanent relationships among phenomena. Examples are organisms changing as they grow, the cycle of seasons, the ebb and flow of tides, cycles of unemployment, weather changes and stock exchange indices. Some of these change processes involve and can be described or modelled by straightforward mathematical functions: linear, exponential, periodic or logistic, either discrete or continuous. But many relationships fall into different categories and data analysis is often essential to determine the kind of relationship that is present. Mathematical relationships often take the shape of equations or inequalities, but relations of a more general nature (*e.g.* equivalence, divisibility, inclusion, to mention but a few) may appear as well.

In order to be sensitive to the patterns of change, Stewart (1990) recommends:

- Representing changes in a comprehensible form
- Understanding the fundamental types of change
- Recognising particular types of change when they occur
- Applying these techniques to the outside world
- Controlling a changing universe to the best advantage

*Change and relationships* can be represented in a variety of ways including numerical (for example in a table), symbolical, graphical, algebraic and geometrical. Translation between these representations is of key importance, as is the recognition of an understanding of fundamental relationships and types of change. Students should be aware of the concepts of linear growth (additive process), exponential growth (multiplicative process) and periodic growth, as well as logistic growth, at least informally as a special case of exponential growth.

Students should also see the relationships among these models – the key differences between linear and exponential processes, the fact that percentage growth is identical with exponential growth, how logistic growth occurs and why, either in continuous or discrete situations.

Changes occur in a system of interrelated objects or phenomena where the elements influence each other. In the examples mentioned in the summary, all phenomena changed over time. But there are many examples in real life of matters in which objects are interrelated in a multitude of ways. For example:

*If the length of the string of a guitar is halved, the new tone is an octave higher than the original tone. The tone is therefore dependent on the string length.*

*When we deposit money into a bank account, we know that the account balance will depend on the size, frequency and number of deposits and withdrawals, and the interest rates.*



Relationships lead to dependency. Dependency concerns the fact that properties and changes of certain mathematical objects may depend on or influence properties and changes of other mathematical objects. Mathematical relationships often take the form of equations or inequalities, but relations of a more general nature may appear as well.

*Change and relationships* involves functional thinking. Functional thinking – that is, thinking in terms of and about relationships – is one of the most fundamental disciplinary aims of the teaching of mathematics (MAA, 1923). For 15-year-old students this includes students having a notion of rate of change, gradients and steepness (although not necessarily in a formal way), and dependence of one variable on another. They should be able to make judgements about how fast processes are taking place, also in a relative way.

This overarching idea closely relates to aspects of other overarching ideas. The study of patterns in numbers can lead to intriguing relationships: the study of Fibonacci numbers and the Golden Ratio are examples. The Golden Ratio is a concept that plays a role in geometry as well. Many more examples of *change and relationships* can be found in *space and shape:* the growth of an area in relation to the growth of a perimeter or diameter. Euclidean geometry lends itself also to the study of relationships. A well-known example is the relationship between the three sides of a triangle. If the length of two sides is known, the third is not determined, but the interval in which it lies is known: the interval's endpoints are the absolute value of the difference between the other two sides, and their sum, respectively. Several other similar relationships exist for the various elements of a triangle.

*Uncertainty* lends itself to various problems that can be viewed from the perspective of *change and relationships*. If two fair dice have been rolled and one of them shows four, what is the chance that the sum exceeds seven? The answer (50%) relies on the dependency of the probability at issue on the set of favourable outcomes. The required probability is the proportion of all such outcomes compared with all possible outcomes, which is a functional dependency.

### Change and relationships examples

## Mathematics Example 5: SCHOOL EXCURSION

A school class wants to rent a coach for an excursion, and three companies are contacted for information about prices.

Company A charges an initial rate of 375 zed plus 0.5 zed per kilometre driven. Company B charges an initial rate of 250 zed plus 0.75 zed per kilometre driven. Company C charges a flat rate of 350 zed up to 200 kilometres, plus 1.02 zed per kilometre beyond 200 km.

---

### Question 1: SCHOOL EXCURSION

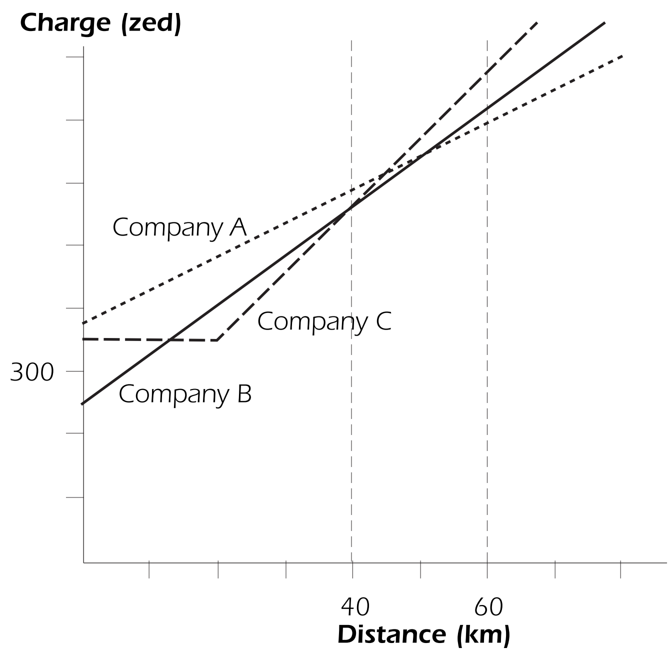
Which company should the class choose, if the excursion involves a total travel distance of somewhere between 400 and 600 km?



Leaving aside the fictitious elements of the context, this problem could conceivably occur. Its solution requires the formulation and activation of several functional relationships, and equations and inequations. It can be dealt with by graphical as well as algebraic means, or combinations of both. The fact that the total travel distance in the excursion is not known exactly also introduces links to the *uncertainty* overarching idea, discussed in a later section.

A graphical representation of the problem is presented below.

Figure 3.6 ■ **Excursion charges for three bus companies**




The following is another example of *change and relationships*.

## Mathematic Example 6: CELL GROWTH

Doctors are monitoring the growth of cells. They are particularly interested in the day that the cell count will reach 60 000 because then they have to start an experiment. The table of results is:

| **Time (days)** | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 |
|---|---|---|---|---|---|---|---|---|---|
| **Cells** | 597 | 893 | 1 339 | 1 995 | 2 976 | 2 976 | 14 719 | 21 956 | 32 763 |

### Question 1: CELL GROWTH

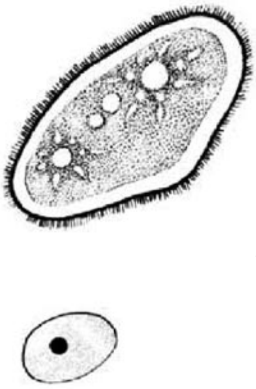
When will the number of cells reach 60 000?

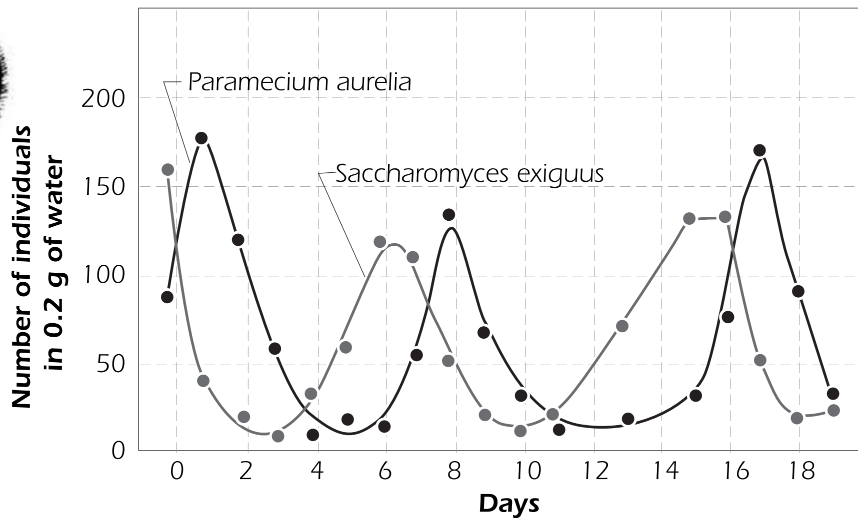


### Mathematic Example 7: PREY-PREDATOR

The following graph shows the growth of two living organisms: the Paramecium and Saccharomyces.:

**Prey Predator Model**






#### Question 1: PREY-PREDATOR

One of the two animals (predator) eats the other one (prey). Looking at the graph, can you judge which one is the prey and which one the predator?

One property of prey-predator phenomena is expressed as: The rate of growth of predators is proportional to the number of available prey. Does this property hold for the above graphs?

#### Quantity

Important aspects of *quantity* include an understanding of relative size, the recognition of numerical patterns, and the use of numbers to represent quantities and quantifiable attributes of real-world objects (counts and measures). Furthermore, *quantity* deals with the processing and understanding of numbers that are represented to us in various ways.

An important aspect of dealing with *quantity* is quantitative reasoning. Essential components of quantitative reasoning are number sense, representing numbers in various ways, understanding the meaning of operations, having a feel for the magnitude of numbers, mathematically elegant computations, mental arithmetic and estimating.

If a magnitude is measured, number uses most important in everyday life are seen: length, area, volume, height, speed, mass, air pressure, money value are all quantified using measures.

Quantitative reasoning is an important aspect of dealing with quantities. It includes:

- Number sense
- Understanding the meaning of operations
- Having a feel for the magnitude of numbers



- Elegant computations
- Mental arithmetic
- Estimations

The "meaning of operations" includes the ability to perform operations involving comparisons, ratios and percentages. Number sense addresses issues of relative size, different representations of numbers, equivalent form of numbers and using understanding of these things to describe attributes of the world.

*Quantity* also includes having a feeling for quantities and estimation. In order to be able to test numerical results for reasonableness, one needs a broad knowledge of quantities (measures) in the real world. Is the average speed of a car 5, 50 or 500 km/h? Is the population of the world 6 million, 600 million, 6 billion, or 60 billion? How tall is a tower? How wide is a river? The ability to make quick order-of-magnitude approximations is of particular importance, especially when viewed in light of the increasing use of electronic calculating tools. One needs to be able to see that 33 x 613 is something around 20 000. To achieve this skill one does not need extensive training in mental execution of traditional written algorithms, but a flexible and smart application of place value understanding and single-digit arithmetic (Fey, 1990).

Using number sense in an appropriate way, students can solve problems requiring direct, inverse and joint proportional reasoning. They are able to estimate rates of change and provide a rationale for the selection of data and level of precision required by operations and models they use. They can examine alternative algorithms, showing why they work or in what cases they fail. They can develop models involving operations, and relationships between operations, for problems involving real-world data and numerical relations requiring operations and comparisons (Dossey, 1997).

In the overarching idea *quantity*, there is a place for 'elegant' quantitative reasoning like that used by Gauss, as discussed in the following example. Creativity coupled with conceptual understanding should be valued at the level of schooling that includes 15-year-old students.

### Quantity examples

## Mathematic Example 8: GAUSS

Karl Friedrich Gauss's (1777-1855) teacher had asked the class to add together all the numbers from 1 to 100. Presumably the teacher's aim was to keep the students occupied for a time. But Gauss was an excellent quantitative reasoner and spotted a shortcut to the solution. His reasoning went like this:

You write down the sum twice, once in ascending order, then in descending order, like this:

1 + 2 + 3 + …………. + 98 + 99 + 100

100 + 99 + 98 +…………+ 3 + 2 + 1

Now you add the two sums, column by column, to give:

101 + 101 +……………..+ 101 + 101

As there are exactly 100 copies of the number 101 in this sum its value is: 100 x 101 = 10 100.

Since this product is twice the answer to the original sum, if you halve it, you obtain the answer: 5 050.



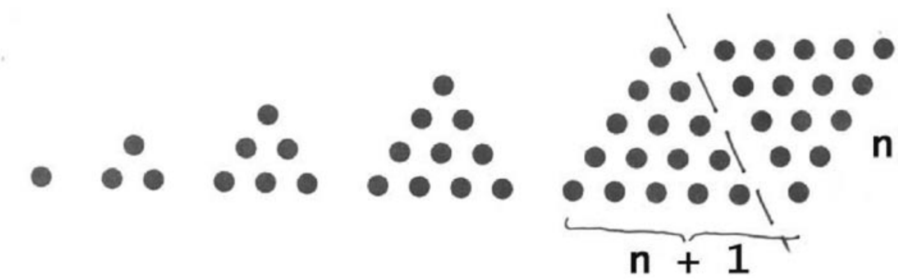
### *Triangular numbers*

This example of quantitative thinking involving patterns of numbers can be taken a little further to demonstrate a link with a geometric representation of that pattern, by showing the formula that gives the general situation for Gauss's problem:

$$1 + 2 + 3 + \ldots + n = n(n + 1)/2$$

This formula also captures a geometric pattern that is well known: numbers of the form n(n+1)/2 are called triangular numbers, since they are exactly the numbers that are obtained by arranging balls in an equilateral triangle. The first five triangular numbers 1, 3, 6, 10 and 15 are shown in Figure 3.7.

Figure 3.7 ■ **The first five triangular numbers**




### *Proportional reasoning*

It will be interesting to see how students in different countries solve problems that lend themselves to the use of a variety of strategies. Differences can be expected especially in the area of proportional reasoning. In certain countries, mainly one strategy per item is likely to be used, while in other countries more strategies will be used. Also, similarities in reasoning will appear in solving problems that do not look very similar. This is in line with recent research results on TIMSS data (Mitchell, J. *et al.,* 2000). The following three items illustrate this point about different strategies and the relationships among them:

1. Tonight you're giving a party. You want to buy 100 cans of soft drink. How many six-can packs are you going to buy?
2. A hang-glider with glide-ratio 1 to 22 starts from a sheer cliff at 120 metres. The pilot is aiming at a spot at a distance of 1 400 metres. Will she reach that spot (under conditions of no wind)?
3. A school wants to rent mini-vans (with seats for eight passengers) for going to a school camp and 98 students need transportation. How many vans does the school need?

The first problem could be seen as a division problem (100+6=) that then leaves the student with an interpretation problem back to the context (what is the meaning of the remainder?). The second problem can be solved by proportional reasoning (for every metre height I can fly a distance of 22 metres, so starting from 120 metres…). The third problem will be solved by many as a division problem. All three problems, however, can be solved using the ratio table method:

**Bottles :**

| | | | | | |
|---|---|---|---|---|---|
| 1 | 10 | 5 | 15 | 2 | 17 |
| 6 | 60 | 30 | 90 | 12 | 102 |

**Flying :**

| | | | |
|---|---|---|---|
| 1 | 100 | 20 | 120 |
| 22 | 2200 | 440 | 2640 |

**Buses :**

| | | | |
|---|---|---|---|
| 1 | 10 | 2 | 13 |
| 8 | 80 | 16 | 104 |

Seeing this similarity is a skill that belongs to *mathematical literacy:* mathematically literate students do not need to look for the one available and appropriate tool or algorithm, but have available to them a wide array of strategies from which they can choose.

## Mathematic Example 9: PERCENTS

Carl went to a store to buy a jacket with a normal price of 50 zed that was on sale for 20% off. In Zedland there is a 5% sales tax. The clerk first added the 5% tax to the price of the jacket and then took 20% off. Carl protested: he wanted the clerk to deduct the 20% discount first and then calculate the 5% tax.

---

### Question 1: PERCENTS

Does it make any difference?

Problems involving this kind of quantitative thinking, and the need to carry out the resulting mental calculations, are encountered frequently when shopping. The ability to effectively handle such problems is fundamental to *mathematical literacy*.

### Uncertainty

Science and technology rarely deal with certainty. Indeed, scientific knowledge is seldom, if ever, absolute – and is even sometimes wrong – so there always remains some uncertainty in even the most scientific predictions. Uncertainty is also present in daily life: uncertain election results, collapsing bridges, stock market crashes, unreliable weather forecasts, poor predictions for population growth or economic models that don't align.

As an overarching idea, *uncertainty* suggests two related topics: data and chance. These phenomena are respectively the subject of mathematical study in statistics and probability. Relatively recent recommendations concerning school curricula are unanimous in suggesting that statistics and probability should occupy a much more prominent place than has been the case in the past (Committee of Inquiry into the Teaching of Mathematics in Schools, 1982; LOGSE, 1990; MSEB, 1990; NCTM, 1989; NCTM, 2000). Specific mathematical concepts and activities that are important in this area are collecting data, data analysis and display/visualisation, probability and inference.

The recommendations on the role of data, statistics and probability in school curricula emphasise data analysis. As a result, it is easy to view statistics in particular as a collection of specific skills. David S. Moore has pointed out what the overarching idea *uncertainty* really is all about. The PISA definition will follow his ideas as presented in *On the Shoulders of Giants* (Steen (1990) and F. James Rutherford's as presented in *Why Numbers Count* (Steen, 1997).



Statistics brings something to mathematics education that is unique and important: reasoning from uncertain empirical data. This kind of statistical thinking should be part of the mental equipment of every intelligent citizen. The core elements are:

- The omnipresence of variation in processes
- The need for data about processes
- The design of data production with variation in mind
- The quantification of variation
- The explanation of variation

Data are not merely numbers, but numbers in a context. Data are obtained by measurement and are represented by a number. Thinking about measurement leads to a mature grasp of why some numbers are informative and others are irrelevant or nonsensical.

The design of sample surveys is a core topic in statistics. Data analysis emphasises understanding the specific data at hand, assuming they represent a larger population. The concept of simple random samples is essential for 15-year-old students to understand the issues related to uncertainty.

Phenomena have uncertain individual outcomes and frequently the pattern of repeated outcomes is random. The concept of probability in the present PISA study will generally be based to situations regarding chance devices like coins, number cubes and spinners, or not too complex real-world situations that can be analysed intuitively, or can feasibly be modelled with these devices.

Uncertainty also appears from sources like natural variation in students' heights, reading scores, incomes of a group of people, etc. A step that is very important, even for 15-year-old students, is to see the study of data and chance as a coherent whole. One such principle is the progression of ideas from simple data analysis to data production to probability to inference.

The important specific mathematical concepts and activities in this area are:

- Producing data
- Data analysis and data display/visualisation
- Probability
- Inference

### Uncertainty examples

The following examples illustrate the *uncertainty* overarching idea.

## Mathematical Example 10: AVERAGE AGE

### Question 1: AVERAGE AGE

If 40% of the population of a country are at least 60 years old, is it then possible for the average age to be 30?



## Mathematics Example 11: GROWING INCOMES?

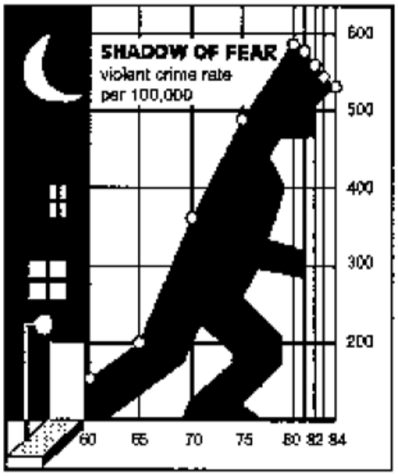
### Question 1: GROWING INCOMES?

Has the income of people in Zedland gone up or down in recent decades? The median money income per household fell: in 1970 it was 34 200 zed, in 1980 it was 30 500 zed and in 1990 31 200 zed. But the income per person increased: in 1970 13 500 zed, in 1980 13 850, and in 1990 15 777 zed.

A household consists of all people living together at the same address. Explain how it is possible for the household income to go down at he same time the per-person income has risen in Zedland.

## Mathematics Example 12: RISING CRIMES

The following graph was taken from the weekly Zedland **News Magazine**:

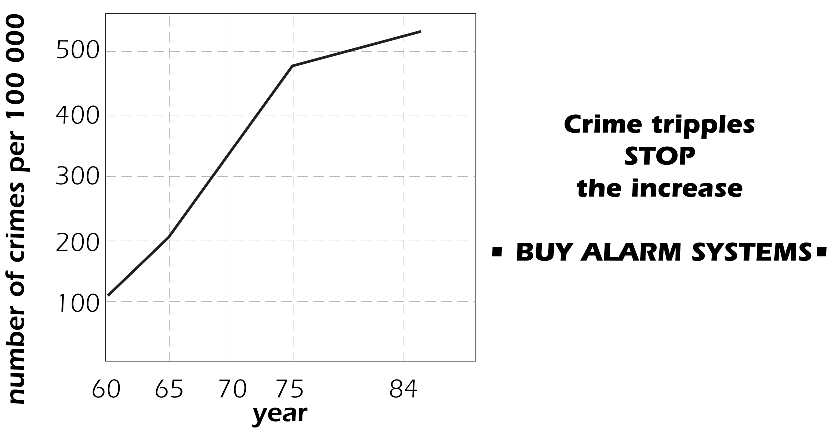



It shows the number of reported crimes per 100 000 inhabitants, starting with five-year intervals, then changing to one-year intervals.

### Question 1: RISING CRIMES

How many reported crimes per 100 000 were there in 1960?

Manufacturers of alarm systems used the same data to produce the following graph:



### Question 2: RISING CRIMES

How did the designers come up with this graph and why?

The police were not too happy with the graph from the alarm systems manufacturers because the police want to show how successful crime fighting has been.

Design a graph to be used by the police to demonstrate that crime has decreased recently.

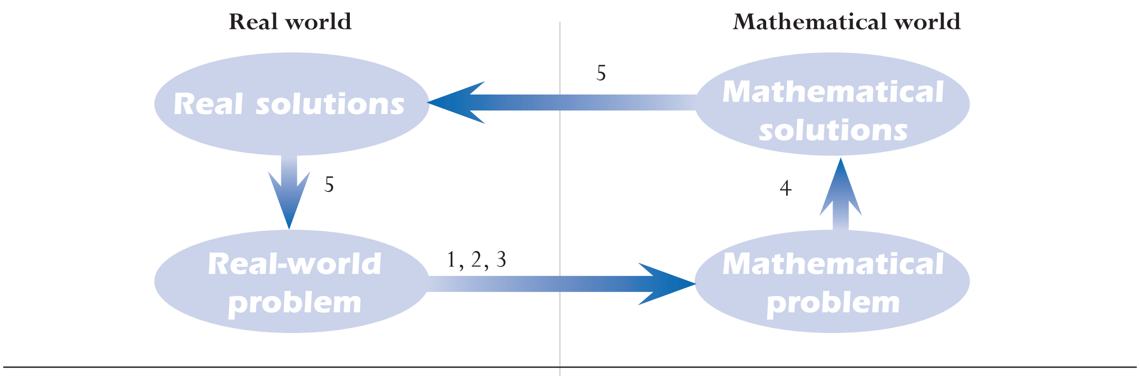
## MATHEMATICAL PROCESSES

### Mathematisation

PISA examines the capacities of students to analyse, reason and communicate mathematical ideas effectively as they pose, formulate, solve and interpret mathematical problems in a variety of situations. Such problem solving requires students to use the skills and competencies they have acquired through schooling and life experiences. In PISA, a fundamental process that students use to solve real-life problems is referred to as mathematisation.

The discussion above of the theoretical basis for the PISA mathematics framework outlined a five-step description of mathematisation. These steps are shown in Figure 3.8 and listed below:

Figure 3.8 ■ **The mathematisation cycle**




- Starting with a problem situated in reality
- Organising it according to mathematical concepts and identifying the relevant mathematics
- Gradually trimming away the reality through processes such as making assumptions, generalising and formalising, which promote the mathematical features of the situation and which transform the real-world problem into a mathematical problem that faithfully represents the situation
- Solving the mathematical problem
- Making sense of the mathematical solution in terms of the real situation, including identifying the limitations of the solution

Mathematisation first involves translating the problem from reality into mathematics. This process includes activities such as:



- Identifying the relevant mathematics with respect to a problem situated in reality
- Representing the problem in a different way, including organising it according to mathematical concepts and making appropriate assumptions
- Understanding the relationships between the language of the problem and the symbolic and formal language needed to understand it mathematically
- Finding regularities, relations and patterns
- Recognising aspects that are isomorphic with known problems
- Translating the problem into mathematics *i.e.* to a mathematical model (de Lange, 1987)

As soon as a student has translated the problem into a mathematical form, the whole process can continue within mathematics. Students pose questions like: "Is there…?", "If so, how many?", "How do I find…?", using known mathematical skills and concepts. They attempt to work on their model of the problem situation, to adjust it, to establish regularities, to identify connections and to create a good mathematical argument. This part of the mathematisation process is generally called the deductive part of the modelling cycle (Blum, 1996; Schupp, 1988). However, processes other than strictly deductive ones may play a part in this stage. This part of the mathematisation process includes:

- Using and switching between different representations
- Using symbolic, formal and technical language and operations
- Refining and adjusting mathematical models, combining and integrating models
- Argumentation
- Generalisation

The last step or steps in solving a problem involve reflecting on the whole mathematisation process and the results. Here students must interpret the results with a critical attitude and validate the whole process. Such reflection takes place at all stages of the process, but it is especially important at the concluding stage. Aspects of this reflecting and validating process are:

- Understanding the extent and limits of mathematical concepts
- Reflecting on mathematical arguments and explaining and justifying results
- Communicating the process and solution
- Critiquing the model and its limits

This stage is indicated in two places in Figure 3.8 by the label "5", where the mathematisation process passes from the mathematical solution to the real solution and where this is related back to the original real-world problem.

### The competencies

The previous section focused on the major concepts and processes involved in mathematisation. An individual who is to engage successfully in mathematisation within a variety of situations, extra- and intra-mathematical contexts, and overarching ideas, needs to possess a number of mathematical competencies which, taken together, can be seen as constituting comprehensive mathematical



competence. Each of these competencies can be possessed at different levels of mastery. Different parts of mathematisation draw differently upon these competencies, both in regard to the particular ones involved and in regard to the required level of mastery. To identify and examine these competencies, PISA has decided to make use of eight characteristic mathematical competencies that rely, in their present form, on the work of Niss (1999) and his Danish colleagues. Similar formulations may be found in the work of many others (as indicated in Neubrand *et al.,* 2001). Some of the terms used, however, have different usage among different authors.

- *Thinking and reasoning:* This involves posing questions characteristic of mathematics ("Is there…?", "If so, how many?", "How do I find…?"); knowing the kinds of answers that mathematics offers to such questions; distinguishing between different kinds of statements (definitions, theorems, conjectures, hypotheses, examples, conditioned assertions); and understanding and handling the extent and limits of given mathematical concepts.
- *Argumentation:* This involves knowing what mathematical proofs are and how they differ from other kinds of mathematical reasoning; following and assessing chains of mathematical arguments of different types; possessing a feel for heuristics ("What can or cannot happen, and why?"); and creating and expressing mathematical arguments.
- *Communication:* This involves expressing oneself, in a variety of ways, on matters with a mathematical content, in oral as well as in written form, and understanding others' written or oral statements about such matters.
- *Modelling:* This involves structuring the field or situation to be modelled; translating reality into mathematical structures; interpreting mathematical models in terms of reality; working with a mathematical model; validating the model; reflecting, analysing and offering a critique of a model and its results; communicating about the model and its results (including the limitations of such results); and monitoring and controlling the modelling process.
- *Problem posing and solving:* This involves posing, formulating and defining different kinds of mathematical problems (for example "pure", "applied", "open ended" and "closed"), and solving different kinds of mathematical problems in a variety of ways.
- *Representation:* This involves decoding and encoding, translating, interpreting and distinguishing between different forms of representation of mathematical objects and situations; the interrelationships between the various representations; and choosing and switching between different forms of representation, according to situation and purpose.
- *Using symbolic, formal and technical language and operations:* This involves decoding and interpreting symbolic and formal language, and understanding its relationship to natural language; translating from natural language to symbolic/formal language; handling statements and expressions containing symbols and formulae; and using variables, solving equations and undertaking calculations.
- *Use of aids and tools:* This involves knowing about, and being able to make use of, various aids and tools (including information technology tools) that may assist mathematical activity and knowing about the limitations of such aids and tools.

PISA does not test the above competencies individually. There is considerable overlap among them, and when using mathematics, it is usually necessary to draw simultaneously on many of the competencies, so that any effort to assess individual ones is likely to result in artificial tasks and unnecessary compartmentalisation of the mathematical literacy domain. The particular competencies



students are able to display vary considerably among individuals. This is partially because all learning occurs through experiences, "with individual knowledge construction occurring through the processes of interaction, negotiation, and collaboration" (de Corte, Greer and Verschaffel, 1996). PISA assumes that much of students' mathematics is learned in schools. Understanding of a domain is acquired gradually. More formal and abstract ways of representing and reasoning emerge over time as a consequence of engagement in activities designed to help informal ideas evolve. *Mathematical literacy* is also acquired through experience involving interactions in a variety of social situations or contexts.

In order to productively describe and report students' capabilities, as well as their strengths and weaknesses from an international perspective, some structure is needed. One way of providing this in a comprehensible and manageable way is to describe clusters of competencies, based on the kinds of cognitive demands needed to solve different mathematical problems.

## COMPETENCY CLUSTERS

PISA has chosen to describe the cognitive activities that these competencies encompass according to three *competency clusters*: the *reproduction* cluster, the *connections* cluster and the *reflection* cluster. In the following sections the three clusters are described and the ways in which the individual competencies are played out in each cluster are discussed.

### The reproduction cluster

The competencies in this cluster essentially involve reproduction of practised knowledge. They include those most commonly used on standardised assessments and classroom tests. These competencies are knowledge of facts and of common problem representations, recognition of equivalents, recollection of familiar mathematical objects and properties, performance of routine procedures, application of standard algorithms and technical skills, manipulation of expressions containing symbols and formulae in standard form, and carrying out computations.

- *Thinking and reasoning:* This involves posing the most basic forms of questions ("How many…?", "How much is…?") and understanding the corresponding kinds of answers ("so many…", "this much…"); distinguishing between definitions and assertions; understanding and handling mathematical concepts in the sorts of contexts in which they were first introduced or have subsequently been practised.
- *Argumentation:* This involves following and justifying standard quantitative processes, including computational processes, statements and results.
- *Communication:* This involves understanding and expressing oneself orally and in writing about simple mathematical matters, such as reproducing the names and the basic properties of familiar objects, citing computations and their results, usually not in more than one way.
- *Modelling:* This involves recognising, recollecting, activating and exploiting well-structured, familiar models; interpreting back and forth between such models (and their results) and reality; and elementary communication about model results.
- *Problem posing and solving:* This involves posing and formulating problems by recognising and reproducing practised standard pure and applied problems in closed form; and solving such problems by invoking and using standard approaches and procedures, typically in one way only.



- *Representation:* This involves decoding, encoding and interpreting familiar, practised standard representations of well known mathematical objects. Switching between representations is involved only when the switching itself is an established part of the representations implied.
- *Using symbolic, formal and technical language and operations:* This involves decoding and interpreting routine basic symbolic and formal language practised in well known contexts and situations; and handling simple statements and expressions containing symbols and formulae, including using variables, solving equations and undertaking calculations by routine procedures.
- *Use of aids and tools:* This involves knowing about and being able to use familiar aids and tools in contexts, situations and ways close to those in which their use was introduced and practised.

Assessment items measuring the *reproduction* cluster competencies can be described with the following key descriptors: reproducing practised material and performing routine operations.

### Examples of reproduction cluster items

## Mathematics Example 13

Solve the equation $7x-3 = 13x+15$

---

## Mathematics Example 14

What is the average of 7, 12, 8, 14, 15, 9?

---

## Mathematics Example 15

1 000 zed is put in a savings account at a bank, with an interest rate of 4%. How many zed will there be in the account after one year?

---

## Mathematics Example 16

In a sprinting event, the 'reaction time' is the time interval between the starter's gun firing and the athlete leaving the starting block. The 'final time' includes both this reaction time, and the running time.

The following table gives the reaction time and the final time of 8 runners in a 100 metre sprint race.

| Lane | Reaction time (sec) | Final time (sec) |
|---|---|---|
| 1 | 0.147 | 10.09 |
| 2 | 0.136 | 9.99 |
| 3 | 0.197 | 9.87 |
| 4 | 0.180 | Did not finish the race |
| 5 | 0.210 | 10.17 |
| 6 | 0.216 | 10.04 |
| 7 | 0.174 | 10.08 |
| 8 | 0.193 | 10.13 |



### Question 1:

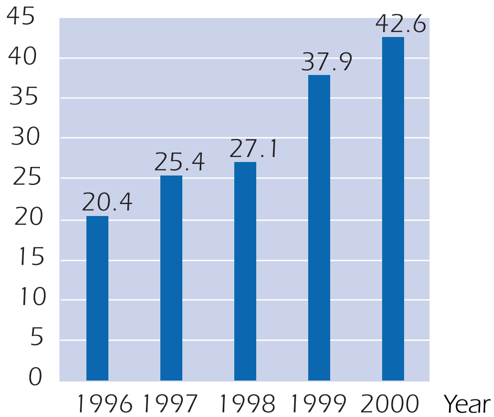
Identify the Gold, Silver and Bronze medallists from this race. Fill in the table below with the medallists' lane number, reaction time and final time.

| Medal | Lane | Reaction time (sec) | Final time (sec) |
|---|---|---|---|
| GOLD | | | |
| SILVER | | | |
| BRONZE | | | |

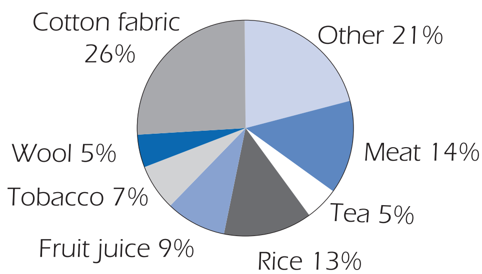
## Mathematics Example 17

The graphics 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:

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.

In order to clarify the boundary for items from the *reproduction* cluster, the *SAVINGS ACCOUNT* problem described in Example 3 provided an example that does not belong to the *reproduction* cluster. This problem takes most students beyond the simple application of a routine procedure and requires the application of a chain of reasoning and a sequence of computational steps that are not characteristic of the *reproduction* cluster competencies.



### The connections cluster

The *connections* cluster competencies build on the *reproduction* cluster competencies in taking problem solving to situations that are not routine, but that still involve familiar or quasi-familiar settings. These competencies include the following:

- *Thinking and reasoning:* This involves posing questions ("How do I find…?", "Which mathematics is involved…?") and understanding the corresponding kinds of answers (provided by means of tables, graphs, algebra, figures, etc.); distinguishing between definitions and assertions and between different kinds of assertions; and understanding and handling mathematical concepts in contexts that are slightly different from those in which they were first introduced or have subsequently been practised.
- *Argumentation:* This involves simple mathematical reasoning without distinguishing between proofs and broader forms of argument and reasoning; following and assessing chains of mathematical arguments of different types; and possessing a feel for heuristics (*e.g.* "What can or cannot happen, or be the case, and why?", "What do I know and what do I want to obtain?").
- *Communication:* This involves understanding and expressing oneself orally and in writing about mathematical matters ranging from reproducing the names and basic properties of familiar objects and explaining computations and their results (usually in more than one way), to explaining matters that include relationships. It also involves understanding others' written or oral statements about such matters.
- *Modelling:* This involves structuring the field or situation to be modelled; translating reality into mathematical structures in contexts that are not too complex but nevertheless different from what students are usually familiar with. It involves also interpreting back and forth between models (and their results) and reality, including aspects of communication about model results.
- *Problem posing and solving:* This involves posing and formulating problems beyond the reproduction of practised standard pure and applied problems in closed form and solving such problems by invoking and using standard approaches and procedures, as well as more independent problem solving processes in which connections are made between different mathematical areas and modes of representation and communication (schemata, tables, graphs, words, pictures).
- *Representation:* This involves decoding, encoding and interpreting familiar and less familiar representations of mathematical objects; choosing and switching between different forms of representation of mathematical objects and situations; and translating and distinguishing between different forms of representation.
- *Using symbolic, formal and technical language and operations:* This involves decoding and interpreting basic symbolic and formal language in less well-known contexts and situations, and handling statements and expressions containing symbols and formulae, including using variables, solving equations and undertaking calculations by familiar procedures.
- *Use of aids and tools:* This involves knowing about and using familiar aids and tools in contexts, situations and ways that are different from those in which their use was introduced and practised.

Items associated with this cluster usually require some evidence of the integration and connection of material from the various overarching ideas, or from different mathematical curriculum strands, or the linking of different representations of a problem.



Assessment items measuring the *connections* cluster of competencies might be described with the following key descriptors: integrating, connecting and modest extension of practised material.

### Examples of connections cluster items

A first example of a *connections* cluster item was given in the *SAVINGS ACCOUNT* problem described in Example 3. Other examples of *connections* cluster items follow.

## Mathematics Example 18: DISTANCE

Mary lives two kilometres from school, Martin five.

---

### Question 1: DISTANCE

How far do Mary and Martin live from each other?

When this problem was originally presented to teachers, many of them rejected it on the ground that it was too easy – one could easily see that the answer is three. Another group of teachers argued that this was not a good item because there was no answer – meaning there is not one single numerical answer. A third reaction was that it was not a good item because there were many possible answers, since without further information the most that can be concluded is that they live somewhere between three and seven kilometres apart, and that is not desirable for an item. A small group thought it was an excellent item, because you have to understand the question, it is real problem solving because there is no strategy known to the student, and it is beautiful mathematics, although you have no clue how students will solve the problem. It is this last interpretation that associates the problem with the *connections* cluster of competencies.

## Mathematics Example 19: THE OFFICE RENTING

The following two advertisements appeared in a daily newspaper in a country where the units of currency are zeds.

| BUILDING A | BUILDING B |
|---|---|
| Office space available<br>58-95 square metres<br>475 **zeds** per month<br>100-120 square metres<br>800 **zeds** per month | Office space available<br>35-260 square metres<br>90 **zeds** per square metre per year |

---

### Question 1: THE OFFICE RENTING

If a company is interested in renting an office of 110 square metres in that country for a year, at which office building, A or B, should the company rent the office in order to get the lower price? Show your work. [© IEA/TIMSS]



### Mathematics Example 20: THE PIZZA

*A pizzeria serves two round pizzas of the same thickness in different sizes. The smaller one has a diameter of 30 cm and costs 30 zeds. The larger one has a diameter of 40 cm and costs 40 zeds. [© PRIM, Stockholm Institute of Education]*

---

#### Question 1: THE PIZZA

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

In both of these problems, students are required to translate a real-world situation into mathematical language, to develop a mathematical model that enables them to make a suitable comparison, to check that the solution fits in with the initial question context and to communicate the result. These are all activities associated with the *connections* cluster.

### The reflection cluster

The competencies in this cluster include an element of reflectiveness on the part of the student about the processes needed or used to solve a problem. They relate to students' abilities to plan solution strategies and implement them in problem settings that contain more elements and may be more 'original' (or unfamiliar) than those in the *connections* cluster. In addition to the competencies described for the *connections* cluster, for the *reflection* cluster the competencies include the following:

- *Thinking and reasoning:* This involves posing questions ("How do I find…?", "Which mathematics are involved…?", "What are the essential aspects of the problem or situation…?") and understanding the corresponding kinds of answers ( provided by tables, graphs, algebra, figures, specification of key points etc.); distinguishing between definitions, theorems, conjectures, hypotheses and assertions about special cases, and reflecting upon or actively articulating these distinctions; understanding and handling mathematical concepts in contexts that are new or complex; and understanding and handling the extent and limits of given mathematical concepts, and generalising results.
- *Argumentation:* This involves simple mathematical reasoning, including distinguishing between proving and proofs and broader forms of argument and reasoning; following, assessing and constructing chains of mathematical arguments of different types; and using heuristics (*e.g.* "What can or cannot happen, or be the case, and why?", "What do I know, and what do I want to obtain?", "Which properties are essential?", "How are the objects related?").
- *Communication:* This involves understanding and expressing oneself orally and in writing about mathematical matters ranging from reproducing the names and basic properties of familiar objects, and explaining computations and their results (usually in more than one way), to explaining matters that include complex relationships, including logical relationships. It also involves understanding others' written or oral statements about such matters.



- *Modelling:* This involves structuring the field or situation to be modelled; translating reality into mathematical structures in contexts that may be complex or largely different from what students are usually familiar with; interpreting back and forth between models (and their results) and reality, including aspects of communication about model results: gathering information and data, monitoring the modelling process and validating the resulting model. It also includes reflecting through analysing, offering a critique, and engaging in more complex communication about models and modelling.
- *Problem posing and solving:* This involves posing and formulating problems well beyond the reproduction of practised standard pure and applied problems in closed form; solving such problems by invoking and using standard approaches and procedures, but also more original problem solving processes in which connections are being made between different mathematical areas and modes of representation and communication (schemata, tables, graphs, words, pictures). It also involves reflecting on strategies and solutions.
- *Representation:* This involves decoding, encoding and interpreting familiar and less familiar representations of mathematical objects; choosing and switching between different forms of representation of mathematical objects and situations, and translating and distinguishing between different forms of representation. It further involves the creative combination of representations and the invention of non-standard ones.
- *Using symbolic, formal and technical language and operations:* This involves decoding and interpreting symbolic and formal language practised in unknown contexts and situations, and handling statements and expressions containing symbols and formulae, including using variables, solving equations and undertaking calculations. It also involves the ability to deal with complex statements and expressions and with unfamiliar symbolic or formal language, and to understand and to translate between such language and natural language.
- *Use of aids and tools:* This involves knowing about and using familiar or unfamiliar aids and tools in contexts, situations and ways that are quite different from those in which their use was introduced and practised. It also involves knowing about limitations of aids and tools.

Assessment items measuring the *reflection* cluster of competencies might be described with the following key descriptors: advanced reasoning, argumentation, abstraction, generalisation and modelling applied to new contexts.

### Examples of reflection cluster items

## Mathematics Example 21: STUDENT HEIGHTS

In a mathematics class one day, the heights of all students were measured. The average height of boys was 160 cm, and the average height of girls was 150 cm. Alena was the tallest – her height was 180 cm. Zdenek was the shortest – his height was 130 cm.

Two students were absent from class that day, but they were in class the next day. Their heights were measured, and the averages were recalculated. Amazingly, the average height of the girls and the average height of the boys did not change.



**Question 1: STUDENT HEIGHT**

Which of the following conclusions can be drawn from this information?

*Circle 'Yes' or 'No' for each conclusion.*

| Conclusion | Can this conclusion be drawn? |
|---|---|
| Both students are girls. | Yes / No |
| One of the students is a boy and the other is a girl. | Yes / No |
| Both students have the same height. | Yes / No |
| The average height of all students did not change. | Yes / No |
| Zdenek is still the shortest. | Yes / No |

The problem is quite complicated in several ways. It requires very precise reading, as superficial reading will likely lead to misinterpretation. Furthermore, it is also difficult to find the crucial mathematical information.

The situation varies within the class and over time. The entity class is used while discussing the average for boys and for girls independently, but subsequently it is stated that Alena is the tallest (girl or student) and Zdenek the shortest (boy or student). If the students do not read carefully they will miss the fact that Alena is a girl and Zdenek a boy.

One obvious difficulty is the fact that the students have to combine the information from the first part of the stimulus (about the different heights) with the second part where the information about two missing students is presented. Here variation over time is seen: two students who were not present in the original setting, but have to be taken into account the next moment in time and so the entity class changes. However, the student solving the problem does not know whether the missing students are boys, girls or one of each. In addition, there is not one problem to solve, but actually five.

Furthermore, to be able to answer correctly the students need to understand mathematically the statistical concepts involved. The problem involves the ability to pose questions ("How do I know…?", "How do I find…?", "What are the possibilities…?" and "What happens if I…?") and the ability to understand and handle the concept of an average in texts that are complex, although the context is familiar.

From this description it will be clear that this item is not only challenging for the students (as shown by the PISA results) but also clearly belongs to the *reflection* cluster.

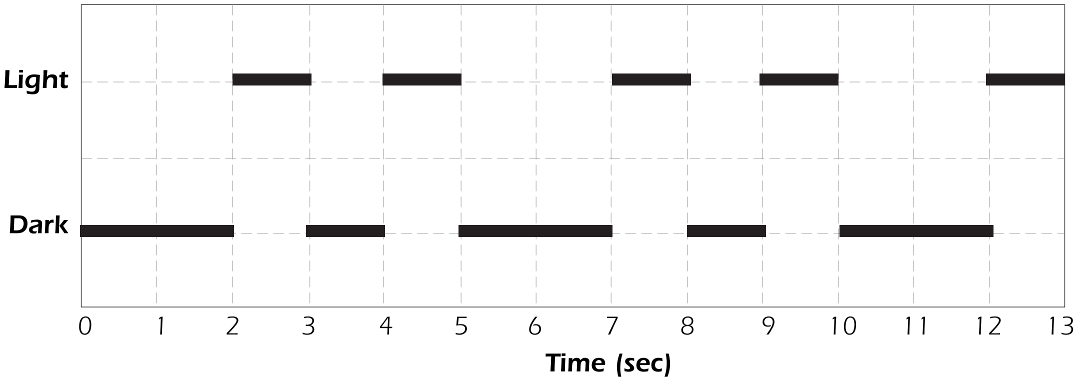


## Mathematics Example 22: LIGHTHOUSE

Lighthouses are towers with a light beacon on top. Lighthouses assist sea ships in finding their way at night when they are sailing close to the shore.

A lighthouse beacon sends out light flashes with a regular fixed pattern. Every lighthouse has its own pattern.

In the diagram below you see the pattern of a certain lighthouse. The light flashes alternate with dark periods.

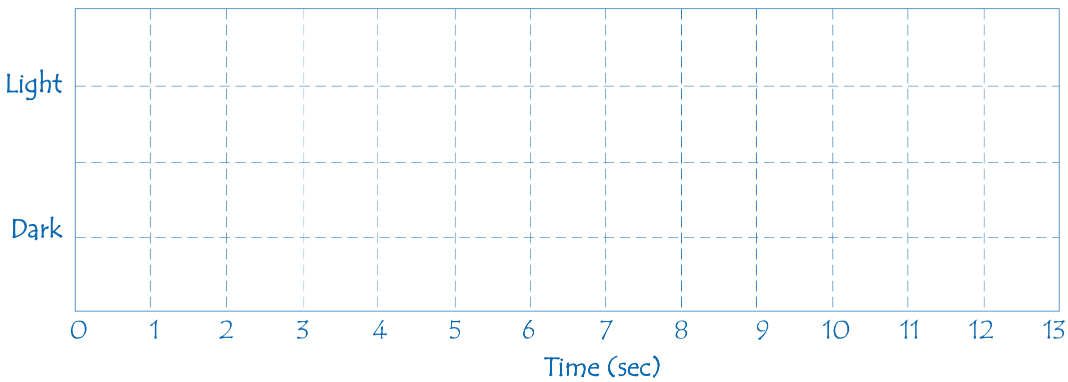



It is a regular pattern. After some time the pattern repeats itself. The time taken by one complete cycle of a pattern, before it starts to repeat, is called the **period.** When you find the period of a pattern, it is easy to extend the diagram for the next seconds or minutes or even hours.

---

### Question 1: LIGHTHOUSE

In the diagram below, make a graph of a possible pattern of light flashes of a lighthouse that sends out light flashes for 30 seconds per minute. The period of this pattern must be equal to 6 seconds.



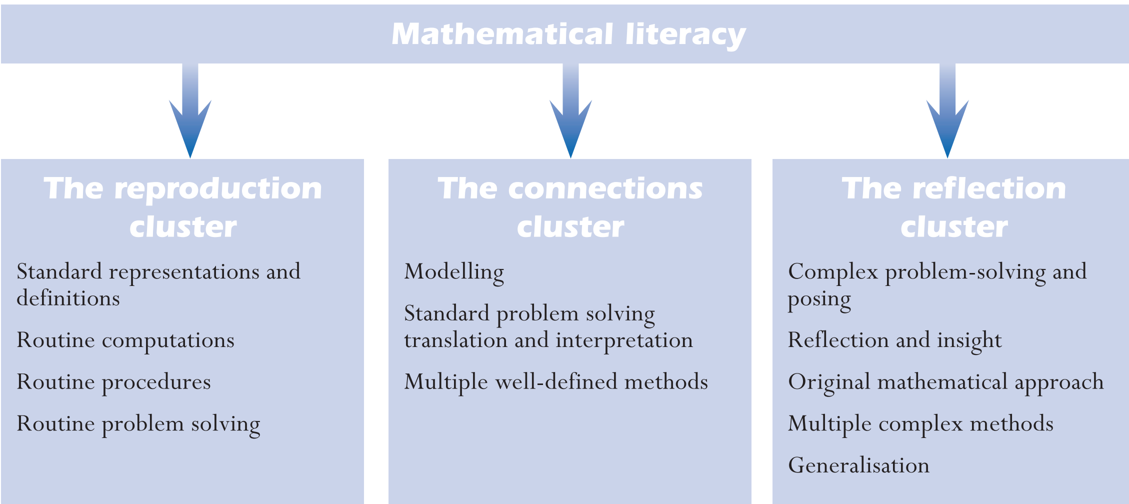

In this example, the students must first understand the introduction in the sense that this kind of graphs is most likely unknown to them, as is the idea of periodicity. In addition, the question posed is of a very open nature: the students are asked to design a possible pattern of light flashes. Many students do not encounter this kind of constructive question at school. However, this constructive aspect is an essential component of being mathematically literate: using mathematical competencies not only in a passive or derived way, but in constructing an answer. Solving the problem demands that two conditions be satisfied: equal amounts of time light and dark (“30 seconds per minute”), and a period of six seconds. This combination makes it essential that the students engage with periodicity at the conceptual level –this involves the *reflection* competency cluster.

In this particular example, the context could be said to favour students living close to an ocean. It should be pointed out, however, that mathematical literacy includes the capacity to use mathematics in contexts different from the local one. This ability to transfer is an essential competency for *mathematical literacy*. While certain students can be in a somewhat favourable position in certain contexts, and others in other contexts the item by country analysis gives no indication that this is the case: landlocked countries did not perform differently from countries bordering on oceans.

### Classification of items by competency cluster

Figure 3.9 summarises the distinctions between the competency clusters.

**Figure 3.9 Diagrammatic representation of the competency clusters**




It is possible to use the competency descriptions in the preceding pages to classify mathematics items and thereby to assign them to one of the competency clusters. One way to do this is to analyse the demands of the item, then to rate each of the eight competencies for that item, according to which of the three clusters provided the most fitting description of item demands in relation to that competency. If any of the competencies were rated as fitting the description for the *reflection* cluster, then the item would be assigned to the *reflection* competency cluster. If not, but one or more of the competencies were rated as fitting the description for the *connections* cluster, then the item would be assigned to that cluster. Otherwise, the item would be assigned to the *reproduction* cluster, since all competencies would have been rated as fitting the competency descriptions for that cluster.

## ASSESSING MATHEMATICAL LITERACY

### Task characteristics

This section considers in more detail features of the assessment tasks that are used to assess students. The nature of the tasks and the item format types are described.



## The nature of PISA mathematics tasks

PISA is an international test of the literacy skills of 15-year-old students. All test items used should be suitable for the population of 15-year-old students in OECD countries.

Trained markers have access to items including some stimulus material or information, an introduction, the actual question and the required solution. In addition, for items with responses that cannot be automatically coded, a detailed coding scheme is developed to enable trained markers across the range of participating countries to code the student responses in a consistent and reliable way.

In an earlier section of this framework, the situations to be used for PISA mathematics items were discussed in some detail. For PISA 2006, each item was set in one of four situation types: personal, educational/occupational, public and scientific. The items selected for the PISA 2006 mathematics instruments represent a spread across these situation types.

In addition, item contexts that can be regarded as authentic are preferred. That is, PISA values most highly tasks that could be encountered in real-world situations, and that have a context for which the use of mathematics to solve the problem would be authentic. Problems with extra-mathematical contexts that influence the solution and its interpretation are preferred as vehicles for assessing *mathematical literacy*.

Items should relate predominantly to the overarching ideas (the phenomenological problem categories) described in the framework. The selection of mathematics test items for PISA 2003 ensured that the four overarching ideas were well represented. Items should embody one or more of the mathematical processes that are described in the framework, and should be identified predominantly with one of the competency clusters.

The level of reading required to successfully engage with an item is considered very carefully in the development and selection of items for inclusion in the PISA test instrument. The wording of items is as simple and direct as possible. Care is also taken to avoid question contexts that would create a cultural bias.

Items selected for inclusion in the PISA test instruments represent a broad range of difficulties, to match the expected wide ability range of students participating in the PISA assessment. In addition, the major classifications of the framework (particularly competency clusters and overarching ideas) should as far as possible be represented with items of a wide range of difficulties. Item difficulties are established in an extensive field trial of test items prior to item selection for the main PISA survey.

## Item types

When assessment instruments are devised, the impact of the item type on student performance, and hence on the definition of the construct that is being assessed, must be carefully considered. This issue is particularly pertinent in a project such as PISA, in which the large-scale cross-national context for testing places serious constraints on the range of feasible item format types.

PISA assesses *mathematical literacy* through a combination of items with open-constructed response types, closed-constructed response types and multiple-choice types. About equal numbers of each of these item format types are used in constructing the test instruments.





Based on experience in developing and using test items for PISA 2000, the multiple-choice type is generally regarded as most suitable for assessing items that would be associated with the *reproduction* and *connections* competency cluster. For an example of this item type, see Example 15, which shows an item that would be associated with the *connections* competency cluster and with a limited number of defined response-options. To solve this problem, students must translate the problem into mathematical terms, devise a model to represent the periodic nature of the context described, and extend the pattern to match the result with one of the given options.

## Mathematics Example 23: SEAL

*A seal has to breathe even if it is asleep. Martin observed a seal for one hour. At the start of his observation the seal dived to the bottom of the sea and started to sleep. In 8 minutes it slowly floated to the surface and took a breath.*

*In 3 minutes it was back at the bottom of the sea again and the whole process started over in a very regular way.*

---

### Question 1: SEAL

After one hour the seal was:

A. At the bottom

B. On its way up

C. Breathing

D. On its way down

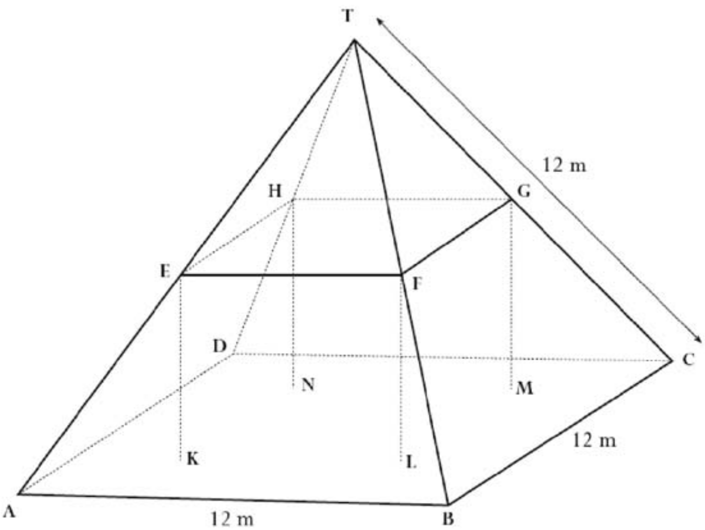
For some of the higher-order goals and more complex processes, other item types are often preferred. Closed-constructed response items pose questions similar to multiple-choice items, but students are asked to produce a response that can be easily judged to be either correct or incorrect. For items in this type, guessing is not likely to be a concern, and the provision of distractors (which influence the construct that is being assessed) is not necessary. For example, for the problem in Example 16 there is one correct answer and many possible incorrect answers.



## Mathematics Example 24: FARMS

Here you see a photograph of a farmhouse with a roof in the shape of a pyramid.

Below is a student's mathematical model of the farmhouse roof with measurements added.




The attic floor, ABCD in the model, is a square. The beams that support the roof are the edges of a block (rectangular prism) EFGHKLMN. E is the middle of AT, F is the middle of BT, G is the middle of CT and H is the middle of DT. All the edges of the pyramid in the model have length 12 m.

---

### Question 1: FARMS

Calculate the area of the attic floor ABCD.

The area of the attic floor ABCD = ........................... m$^2$

Open-constructed response items require a more extended response from the student, and the process of producing a response frequently involves higher-order cognitive activities. Often such items not only ask the student to produce a response, but also require the student to show



the steps taken or to explain how the answer was reached. The key feature of open-constructed response items is that they allow students to demonstrate their abilities by providing solutions at a range of levels of mathematical complexity, exemplified in Example 17.

## Mathematics Example 25: INDONESIA

Indonesia lies between Malaysia and Australia. Some data of the population of Indonesia and its distribution over the islands is shown in the following table:

| Region | Surface area (Km2) | Percentage of total area | Population in 1980 (millions) | Percentage of total population |
|---|---|---|---|---|
| Java/Madura | 132 187 | 6.95 | 91 281 | 61.87 |
| Sumatra | 473 606 | 24.86 | 27 981 | 18.99 |
| Kalimantan (Borneo) | 539 460 | 28.32 | 6 721 | 4.56 |
| Sulawesi (Celebes) | 189 216 | 9.93 | 10 377 | 7.04 |
| Bali | 5 561 | 0.30 | 2 470 | 1.68 |
| Irian Jaya | 421 981 | 22.16 | 1 145 | 5.02 |
| **TOTAL** | **1 905 569** | **100.00** | **147 384** | **100.00** |

One of the main challenges for Indonesia is the uneven distribution of the population over the islands. From the table we can see that Java, which has less than 7% of the total area, has almost 62% of the population.

**Source:** de Lange and Verhage (1992). Used with permission

### Question 1: INDONESIA

*Design a graph (or graphs) that shows the uneven distribution of the Indonesian population.*

For PISA, about one-third of the mathematics items are open-constructed response items. The responses to these items require coding by trained people who implement a coding rubric that may require an element of professional judgement. Because of the potential for disagreement between markers of these items, PISA implements marker reliability studies to monitor the extent of disagreement. Experience in these types of studies shows that clear coding rubrics can be developed and reliable scores can be obtained.

PISA makes some use of a unit format in which several items are linked to common stimulus material. Tasks of this format give students the opportunity to become involved with a context or problem by asking a series of questions of increasing complexity. The first few questions are typically multiple-choice or closed-constructed response items, while subsequent ones are typically open-constructed response items. This format can be used to assess each of the competency clusters.

One reason for the use of common stimulus task formats is that it allows realistic tasks to be devised and the complexity of real-life situations to be reflected in them. Another reason relates to the efficient use of testing time, cutting down on the time required for a student to understand the problem. The need to make each scored point independent of others within the task is recognised and taken into account in the design of the PISA tasks and of the response coding and scoring rubrics. The importance of minimising bias that may result from the use of fewer situations is also recognised.



### Assessment structure

The PISA 2003 test instruments, when mathematics was the major PISA assessment domain, contained a total of 210 minutes of testing time. The selected test items were arranged in seven clusters of items, with each item-cluster representing 30 minutes of testing time. The item-clusters were placed in test booklets according to a rotated test design. For the 2006 test cycle, less time is devoted to testing of mathematics, but the item clusters allocated to mathematics are constructed and rotated in a similar way.

The total testing time for mathematics is distributed as evenly as possible across the four overarching ideas (*space and shape, change and relationships, quantity* and *uncertainty*), and the four situations described in the framework (*personal*, *educational/occupational*, *public* and *scientific*). The proportion of items reflecting the three competency clusters (*reproduction*, *connections* and *reflection*) is about 1:2:1. About one-third of the items is in multiple-choice response type, about one-third in closed-constructed response type, and about one-third in open-constructed response type.

### Reporting mathematical proficiency

To summarise data from responses to the PISA test instruments, a five-level described performance scale was created (Masters and Forster, 1996; Masters, Adams, and Wilson, 1999). The scale was created statistically, using an item response modelling approach to scaling ordered outcome data. The overall scale is used to describe the nature of performance by classifying the student performances of different countries in terms of the five described performance levels, and thus provide a frame of reference for international comparisons.

Consideration was given to developing a number of separate reporting scales. Such subscales could most obviously be based on the three competency clusters, or on the four overarching ideas. Decisions about the development of separate reporting scales were made on a variety of grounds, including psychometric considerations, following analysis of the data generated by the PISA assessments. To facilitate these possibilities, it was necessary to ensure that sufficient items were selected for inclusion in the PISA test instrument from each potential reporting category. Moreover, items within each such category needed to have a suitably wide range of difficulties.

The competency clusters described earlier in this framework reflect conceptual categories of broadly increasing cognitive demand and complexity, but do not strictly reflect a hierarchy of student performances based on item difficulty. Conceptual complexity is only one component of item difficulty that influences levels of performance. Others include familiarity, recent opportunity to learn and practice, etc. Thus, a multiple-choice item involving competencies from the *reproduction* cluster (for example, "Which of the following is a rectangular parallelepiped?" followed by pictures of a ball, a can, a box, and a square) may be very easy for students who have been taught the meaning of these terms, but very difficult for others because of lack of familiarity with the terminology used. While it is possible to imagine relatively difficult *reproduction* cluster items and relatively easy *reflection* cluster items, and as far as possible items with a range of difficulties within each cluster type should be included, one would expect a broadly positive relationship between competency clusters and item difficulty.

Factors that underpin increasing levels of item difficulty and mathematical proficiency include the following:



- *The kind and degree of interpretation and reflection required:* This includes the nature of demands arising from the problem context; the extent to which the mathematical demands of the problem are apparent or to which students must impose their own mathematical construction on the problem, and the extent to which insight, complex reasoning and generalisation are required.
- *The kind of representation skills required:* These include problems where only one mode of representation is used as well as problems where students have to switch between different modes of representation or to find appropriate modes of representation themselves.
- *The kind and level of mathematical skill required:* These include single-step problems requiring students to reproduce basic mathematical facts and perform simple computation processes, multi-step problems involving more advanced mathematical knowledge, complex decision-making, information processing, and problem solving and modelling skills.
- *The kind and degree of mathematical argumentation required:* These include problems where no arguing is necessary at all, problems where students may apply well-known arguments, and problems where students have to create mathematical arguments or to understand other people's argumentation or judge the correctness of given arguments or proofs.

At the lowest described proficiency level, students typically carry out single-step processes that involve recognition of familiar contexts and mathematically well-formulated problems, reproducing well-known mathematical facts or processes, and applying simple computational skills.

At higher proficiency levels, students typically carry out more complex tasks involving more than a single processing step. They also combine different pieces of information or interpret different representations of mathematical concepts or information, recognising which elements are relevant and important and how they relate to one another. They typically work with given mathematical models or formulations, which are frequently in algebraic form, to identify solutions, or they carry out a small sequence of processing or calculation steps to produce a solution.

At the highest proficiency level, students take a more creative and active role in their approach to mathematical problems. They typically interpret more complex information and negotiate a number of processing steps. They produce a formulation of a problem and often develop a suitable model that facilitates its solution. Students at this level typically identify and apply relevant tools and knowledge in an unfamiliar problem context. They likewise demonstrate insight in identifying a suitable solution strategy, and display other higher-order cognitive processes such as generalisation, reasoning and argumentation to explain or communicate results.

### Aids and tools

PISA policy allows students to use calculators and other tools as they are normally used in school.

This represents the most authentic assessment of what students can achieve, and provides the most informative comparison of the performance of education systems. A system's choice to allow students to access and use calculators is no different, in principle, from other instructional policy decisions that are made by systems and are not controlled by PISA.

Students who are used to having a calculator available to assist them in answering questions would be disadvantaged if this resource were taken away.



## SUMMARY

The aim of the PISA study, with regard to mathematics, is to develop indicators that show how effectively countries have prepared their 15-year-old students to become active, reflective and intelligent citizens from the perspective of their uses of mathematics. To achieve this, PISA has developed assessments that focus on determining the extent to which students can use what they have learned. They emphasise mathematical knowledge and understanding to solve problems that arise out of day-to-day experience and provide a variety of problems with varying degrees of built-in guidance and structure, but pushing towards authentic problems where students must do the thinking themselves.

This framework describes and illustrates PISA 2006 definition of *mathematical literacy* and sets the context for the items. The major components of the mathematics framework, consistent with the other PISA frameworks, include contexts for the use of mathematics, mathematical content and mathematical processes, each of which flows directly out of the literacy definition. The discussions of context and content emphasise features of the problems that confront students as citizens, while the discussions of processes emphasise the competencies that students bring to bear to solve those problems. These competencies have been grouped into three competency clusters to facilitate a rational treatment of the way complex cognitive processes are addressed within a structured assessment program.



# References


**Baumert, J.** and **O. Köller** (1998), "Interest Research in Secondary Level I: An Overview" in L. Hoffmann, A. Krapp, K.A. Renniger and J. Baumert (eds.), *Interest and Learning,* Institute for Science Education at the University of Kiel, Kiel.

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

**Blum, W.** (1996), "Anwendungsorientierter Mathematikunterricht – Trends und Perspektiven", in G. Kadunz *et al.* (eds.), *Trends und Perspektiven: Schriftenreihe Didaktik der Mathematik,* Vol. 23, Hoelder-Pichler-Tempsky, Wein.

**Bogner, F.** and **M. Wiseman** (1999), "Toward Measuring Adolescent Environmental Perception", European Psychologist 4 (3).

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

**Bybee, R.** (1997b), "Toward an Understanding of Scientific Literacy" in W. Gräber and C. Bolte (eds.), *Scientific Literacy: An International Symposium,* Institute for Science Education at the University of Kiel, Kiel.

**Committee of Inquiry into the Teaching of Mathematics in Schools** (1982), *Mathematics Counts* (The Cockcroft Report), Her Majesty's Stationery Office, London.

**de Corte, E., B. Greer** and **L. Verschaffel** (1996). "Mathematics Teaching and Learning", in D. C Berliner and R. C. Calfee (eds.), *Handbook of Educational Psychology,* Macmillan, New York.

**de Lange, J**. (1987), *Mathematics, Insight and Meaning,* CD-Press, Utrecht.

**de Lange, J.** (1995), "Assessment: No Change without Problems", in T.A. Romberg (ed.), *Reform in School Mathematics,* SUNY Press, Albany.

**de Lange, J.** and **H. Verhage** (1992), *Data Visualization,* Sunburst, Pleasantville.

**Dossey, J.A**. (1997), "Defining and Measuring Quantitative Literacy", in L. A. Steen, (ed.), *Why Numbers Count,* The College Board, New York.

**Eagles, P.F.J.** and **R. Demare** (1999), "Factors Influencing Children's Environmental Attitudes", *The Journal of Environmental Education,* 30 (4).

**Fensham, P.J.** (1985), "Science for All: A Reflective Essay", *Journal of Curriculum Studies* 17 (4).

**Fensham, P.J.** (2000), "Time to Change Drivers for Scientific Literacy", *Canadian Journal of Science, Mathematics, and Technology Education* 2, 9-24.

**Fey, J.** (1990), "Quantity", in L.A. Steen (ed.), *On the Shoulders of Giants: New Approaches to Numeracy,* National Academy Press, Washington, D.C.

**Fleming, R.** (1989), "Literacy for a Technological Age", *Science Education* 73 (4).

**Freudenthal, H.** (1973), *Mathematics as an Educational Task*, D. Reidel, Dordrecht.

**Freudenthal, H.** (1983), *Didactical Phenomenology of Mathematical Structures*, D. Reidel, Dordrecht.

**Gardner, P.L.** (1975), "Attitudes to Science: A Review", *Studies in Science Education* 2.

**Gardner, P. L.** (1984), "Students' Interest in Science and Technology: An International Overview" in M. Lehrke, L. Hoffmann and P. L. Gardner (eds.), *Interests in Science and Technology Education,* Institute for Science Education at the University of Kiel, Kiel.



**Gauld, C.** and **A.A. Hukins** (1980), "Scientific Attitudes: A Review", *Studies in Science Education* 7.

**Gee, J.** (1998), *Preamble to a Literacy Program,* Department of Curriculum and Instruction, Madison.

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

**Grünbaum, B.** (1985), "Geometry Strikes Again", *Mathematics Magazine*, 58 (1).

**Hershkowitz, R., B. Parzysz and J. van Dormolen** (1996), "Space and Shape", in Bishop, A. J., K. Clements, C. Keitel, J. Kilpatrick and C. Laborde (eds.), *International Handbook of Mathematics Education, Part 1,* Kluwer Academic Publishers, Dordrecht.

**Kirsch, I.S.** and **P.B. Mosenthal** (1989-1991), "Understanding Documents: A Monthly Column", *Journal of Reading,* International Reading Association, Newark.

**Klopfer, L.** (1971), "Evaluation of Learning in Science", in B. Bloom, J. Hastings and G. Madaus (eds.), *Handbook of Summative and Formative Evaluation of Student Learning,* McGraw-Hill, New York.

**Klopfer, L.E.** (1976), "A Structure for the Affective Domain in Relation to Science Education", *Science Education* 60.

**Koballa, T., A. Kemp** and **R. Evans** (1997), "The Spectrum of Scientific Literacy", *The Science Teacher* 64 (7).

**Kuhn, D.** (1992), "Thinking as Argument", *Harvard Educational Review* 62 (2).

**LaForgia, J.** (1988), "The Affective Domain Related to Science Education and Its Evaluation", *Science Education* 72 (4).

**Langer, J.** (1995), *Envisioning Literature,* International Reading Association, Newark.

**Law, N.** (2002), "Scientific Literacy: Charting the Terrains of a Multifaceted Enterprise", *Canadian Journal of Science, Mathematics, and Technology Education* 2, 151–176.

**LOGSE** (1990), *Ley de Ordenacion General del Sistema Educativo*, LOGSE, Madrid.

**Masters, G., R. Adams** and **M. Wilson** (1999), "Charting Student Progress", in G. Masters and J. Keeves (eds.), *Advances in Measurement in Educational Research and Assessment,* Amsterdam, Elsevier Science.

**Masters, G.** and **M. Forster** (1996), *Progress Maps,* Australian Council for Educational Research, Melbourne.

**Mathematical Association of America - MAA** (1923), *The Re-organization of Mathematics in Secondary Education: A Report of the National Committee on Mathematical Requirements,* MAA, Oberlin.

**Mathematical Sciences Education Board - MSEB** (1990), *Reshaping School Mathematics: A Philosophy and Framework of Curriculum,* National Academy Press, Washington, D.C.

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

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

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

**Mitchell, J., E. Hawkins, P. Jakwerth, F. Stancavage** and **J. Dossey** (2000), *Student Work and Teacher Practice in Mathematics,* National Center for Education Statistics, Washington, D.C.

**National Council of Teachers of Mathematics - NCTM** (1989), *Curriculum and Evaluation Standards for School Mathematics,* NCTM, Reston.



**NCTM** (2000), *Principles and Standards for School Mathematics,* NCTM, Reston.

**Neubrand, M., R. Biehler, W. Blum, E. Cohors-Fresenborg, L. Flade, N. Knoche, D. Lind, W. Löding, G. Möller** and **A. Wynands** (Deutsche OECD/PISA-Expertengruppe Mathematik) (2001), "Grundlagen der Ergänzung des internationalen OECD/PISA-Mathematik-Tests in der deutschen Zusatzerhebung", *Zentralblatt für Didaktik der Mathematik* 33 (2).

**Niss, M.** (1999), "Kompetencer og uddannelsesbeskrivelse" *Uddanneise,* 9.

**Norris, S.** and **L. Phillips** (2003), "How Literacy in Its Fundamental Sense is Central to Scientific Literacy", *Science Education* 87 (2).

**OECD** (Organisation for Economic Co-operation and Development) (1999), *Measuring Student Knowledge and Skills: A New Framework for Assessment,* OECD, Paris.

**OECD** (2000), *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical, and Scientific Literacy,* OECD, Paris.

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

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

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

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

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

**OECD** (2005), *Are Students Ready for a Technology-Rich World? What PISA Studies Tell Us,* OECD, Paris.

**Osborne, J., S. Erduran, S. Simon** and **M. Monk** (2001), "Enhancing the Quality of Argumentation in School Science", *School Science Review* 82 (301).

**Osborne, J., S. Simon** and **S. Collins** (2003), "Attitudes towards Science: A Review of the Literature and its Implications", *International Journal of Science Education* 25 (9).

**Rickinson, M.** (2001), "Learners and Learning in Environmental Education: A Critical Review of the Evidence", *Environmental Education Research* 7 (3).

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

**Schibeci, R.A.** (1984), "Attitudes to Science: An Update", *Studies in Science Education* 11.

**Schupp, H.** (1988), "Anwendungsorientierter mathematikunterrricht in der sekundarstufe I zwischen tradition und neuen impulsen", *Der Mathematikunterricht* 34 (6).

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

**Steen, L.A**. (ed.) (1997), *Why Numbers Count: Quantitative Literacy for Tomorrow's America,* The College Board, New York.

**Stewart, K.** (1990), "Change", in L. A. Steen (ed.), *On the Shoulders of Giants: New Approaches to Numeracy,* National Academy Press, Washington, D.C.

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

**Stiggins, R.J.** (1982), "An Analysis of the Dimensions of Job-Related Reading", *Reading World, 82*.




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

**UNESCO** (2003), "UNESCO and the International Decade of Education for Sustainable Development (2005–2015)", *UNESCO International Science, Technology & Environmental Education Newsletter,* Vol. XXVIII, no. 1–2, UNESCO, Paris.

**UNESCO** (2005), *International Implementation Scheme for the UN Decade of Education for Sustainable Development,* UNESCO, Paris.

**Weaver, A.** (2002), "Determinants of Environmental Attitudes: A Five-Country Comparison", *International Journal of Sociology* 32 (1).

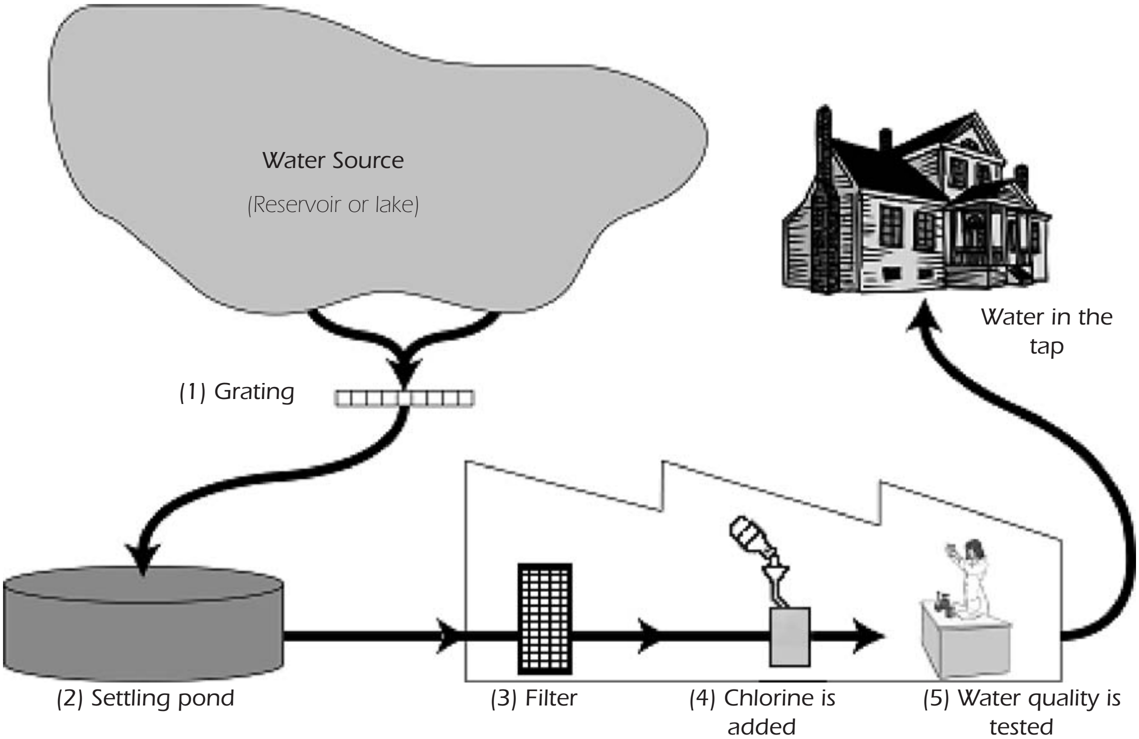
# Annex A

## ADDITIONAL SCIENCE UNITS



## SCIENCE UNIT 1

### Fit for Drinking




The figure above shows how water supplied to houses in cities is made fit for drinking.

---

### Question 1.1

*It is important to have a source of good drinking water. Water found underground is referred to as ground water.*

*Give one reason why there is less bacteria and particle pollution in ground water than in water from surface sources such as lakes and rivers.*

#### Scoring and Comments on Question 1.1

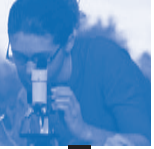
**Full Credit**

Code 11: Responses referring to ground water being filtered through the ground

- *When it goes through sand and dust the water is cleaned.*
- *It has been naturally filtered.*
- *Because when water goes down through the ground it will be strained by rocks and sand.*

Code 12: Responses referring to the ground water being encapsulated and therefore protected from possible pollution; OR that surface water is more easily polluted

- *Ground water is inside the earth and therefore air pollution cannot make it dirty.*
- *Because ground water isn't open, it is located under something.*
- *Lake and rivers can be polluted by the air and you can swim in it and so on, that's why it is not clean.*
- *Because lakes and rivers are polluted by people and animals.*





Code 13: Other correct responses

- Ground water is water without much food for bacteria so they will not survive there.
- Ground water is not in the Sun. There is blue-green algae.

**No Credit**

Code 01: Responses referring to ground water being very clean (information already given)

- Because it has been cleaned.
- Because there is rubbish in lakes and rivers. [Does not explain why.]
- Because there is less bacteria.

Code 02: Responses obviously referring to the cleaning process provided in the figure given in the stimulus

- Because ground water passes through a filter and chlorine is added.
- The ground water passes through a filter that cleans it absolutely.

Code 03: Other responses

- Because it's always moving.
- Because it is not stirred and therefore doesn't bring mud from the bottom.

Code 99: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Earth and space systems (Knowledge of science)

**Application area:** Natural resources

**Setting:** Global

This item focuses attention on two aspects of water quality – particle and bacterial contamination. Answering it requires the application of scientific knowledge to explain why untreated ground water is less contaminated than untreated surface water.

The availability of clean drinking water is of everyday importance for people everywhere, although as an issue its prominence will differ with circumstance. Item classification is consistent with the necessity to draw on knowledge of science in explaining a phenomenon.

This item performed well in the trial with good discrimination and was of moderate difficulty being correctly answered by about two-thirds of students.

---

### Question 1.2

The cleaning of water often happens in several steps, involving different techniques. The cleaning process shown in the figure involves four steps (numbered 1–4). In the second step, the water is collected in a settling pond.

In what way does this step make the water cleaner?

A. The bacteria in the water die.

B. Oxygen is added to the water.

C. Gravel and sand sink to the bottom.

D. Toxic substances are broken down.



### Scoring and Comments on Question 1.2

**Full Credit**

Code 1: C. Gravel and sand sink to the bottom.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Health

**Setting:** Social

The unit context identifies the ways water collected in reservoirs and lakes is cleaned before being distributed to households. This item involves recognising, or deducing, the purpose of the settling pond. Thus it assesses students' knowledge of sedimentation as a gravitational effect on the particles in the water.

The field trial showed the item to be of average difficulty. It showed good discrimination, although the second option (B) was a weak distractor.

---

### Question 1.3

*In the fourth step of the cleaning process, chlorine is added to the water.*

*Why is chlorine added to the water?*

### Scoring and Comments on Question 1.3

**Full Credit**

Code 1: Responses referring to removing, killing or breaking down bacteria (or microbes or viruses or germs)

- To make it free from bacteria.
- Chlorine kills bacteria.
- To kill all the algae.

**No Credit**

Code 0: Other responses

- The water gets less acid and there will be no algae.
- It is like fluoride.
- To clean out the water a bit more and kill the left over things. ["Things" is not specific enough.]
- To keep it clean and drinkable.

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically



**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Social

As with the previous item, the context has everyday relevance in that citizens should know something of the treatments applied to their drinking water.

Knowledge of the effects of chlorine on living organisms is applied in explaining why chlorine is added to the water. The knowledge category is therefore "Living systems".

The item performed reasonably well in the field trial with adequate discrimination. Overall, it was of low–medium difficulty but it was significantly more difficult than this in a small number of countries.

---

## Question 1.4

Suppose that the scientists involved in the testing of water at the water plant discover that there are some dangerous bacteria in the water *after* the cleaning process is completed.

What should people at home do with this water before drinking it?

### Scoring and Comments on Question 1.4

#### Full Credit

Code 11: Responses referring to boiling the water

- Boil it.

Code 12: Responses referring to other methods of cleaning that are possible to do safely at home

- Treat the water with chlorine tablets (e.g. Puratabs).
- Use a micropore filter.

#### No Credit

Code 01: Responses referring to "professional" methods of cleaning that are impossible to carry out safely at home, or impractical to carry out at home

- Mix it with chloride in a bucket and then drink it.
- More chloride, chemicals and biological devices.
- Distil the water.

Code 02: Other responses

- Purify it again.
- Use a coffee filter.
- Buy bottled water until the cleaning process is fixed. [Avoids the question being asked.]

Code 99: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Social



This item requires students to know how bacteria can be killed or removed from water by methods that are practical to perform at home. The knowledge category is therefore "Living systems".

The item showed adequate discrimination in the field trial and on average over countries was of low-moderate difficulty. However, there was a large variation in difficulties between groups of countries and so the item was not considered for inclusion in the main study.

---

### Question 1.5

Can drinking polluted water cause the following health problems?

*Circle "Yes" or "No" in each case.*

| Can drinking polluted water cause this health problem? | Yes or No? |
|---|---|
| Diabetes | Yes / No |
| Diarrhoea | Yes / No |
| HIV / AIDS | Yes / No |

### Scoring and Comments on Question 1.5

#### Full Credit

Code 1: All three correct: No, Yes, No, in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal

This item assesses knowledge of whether some common illnesses (genetic, bacterial, viral) can be transmitted in water.

While classified as "Explaining phenomena scientifically", it is a low-level form of this competency in that it can be answered by simple recall of knowledge. The category of knowledge involved is clearly "Living systems".

The item was of low difficulty and showed adequate discrimination. Females were more likely to answer it correctly than were males.

---



**Question 1.6**

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing how water is tested for bacterial contamination | ☐1 | ☐2 | ☐3 | ☐4 |
| b) | Learning more about the chemical treatment of water supplies | ☐1 | ☐2 | ☐3 | ☐4 |
| c) | Learning which diseases are transmitted in drinking water | ☐1 | ☐2 | ☐3 | ☐4 |

**Item type:** Attitudinal

**Attitude:** Interest in learning about science

This item was designed to assess students' interest in learning more about scientific issues related to drinking water. It comes at the end of the unit so that the students are familiar with the context prior to being asked their opinions.

Interest is shown by the students providing an indication of the degree to which they want to acquire information about three separate aspects of the contamination of water and its treatment for drinking purposes.

Exploratory factor analysis of the field trial results showed that all three statements loaded significantly on an "interest" dimension. Very strong interest was shown in learning about which diseases are transmitted in drinking water (statement c) – as was the case with most statements that concerned personal health and safety.

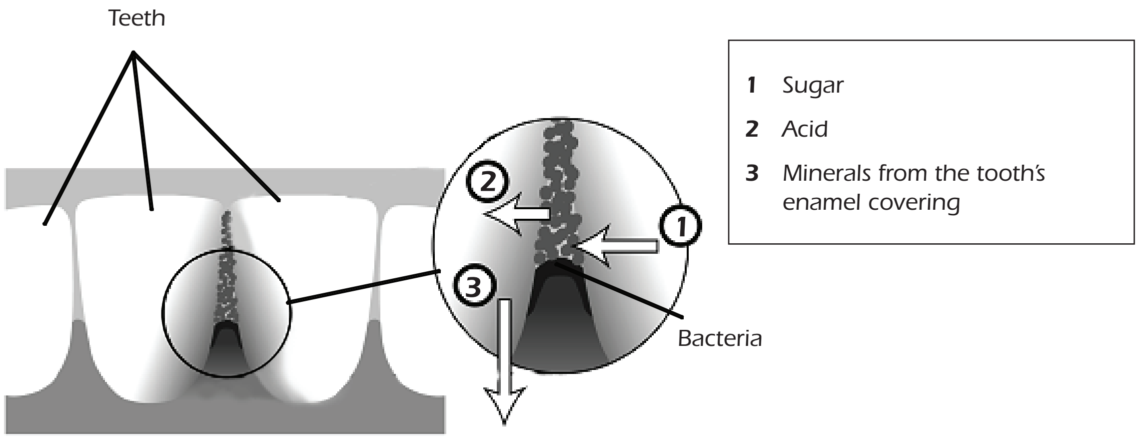
# SCIENCE UNIT 2

## Tooth Decay

Bacteria that live in our mouths cause dental caries (tooth decay). Caries have been a problem since the 1700s when sugar became available from the expanding sugar cane industry.

Today, we know a lot about caries. For example:

- Bacteria that cause caries feed on sugar.
- The sugar is transformed to acid.
- Acid damages the surface of teeth.
- Brushing teeth helps to prevent caries.




### Question 2.1

What is the role of bacteria in dental caries?

A. Bacteria produce enamel.

B. Bacteria produce sugar.

C. Bacteria produce minerals.

D. Bacteria produce acid.

#### Scoring and Comments on Question 2.1

**Full Credit**

Code 1: D. Bacteria produce acid.

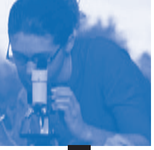
**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Using scientific evidence





**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Health

**Setting:** Personal

Information about tooth decay is provided using a diagram and associated text. Students are required to select the conclusion that follows from the given information.

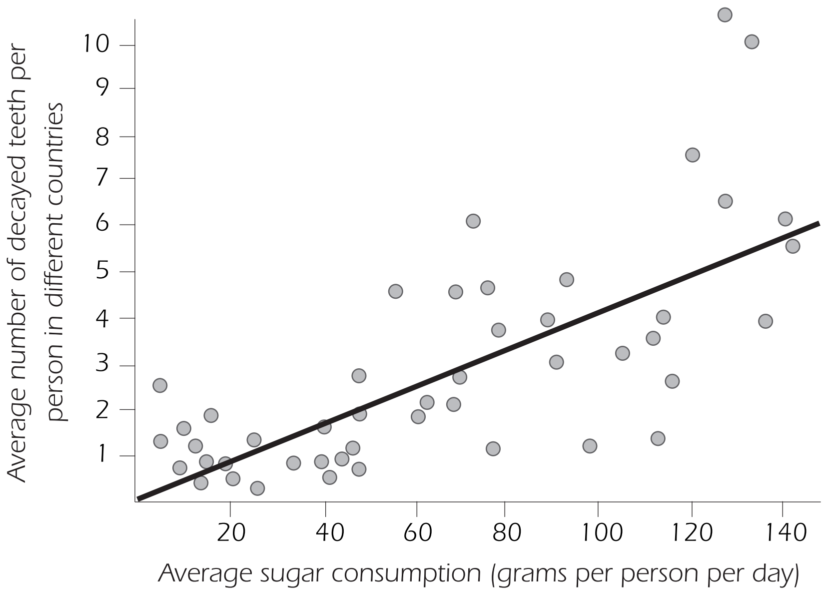
Because the science involved concerns only the capacity to utilise evidence in drawing a conclusion, the item mainly assesses "Knowledge about science".

This item performed well in the field trial showing good discrimination and was found to be relatively easy.

---

## Question 2.2

The following graph shows the consumption of sugar and the amount of caries in different countries.




Each country is represented by a dot in the graph.

Which one of the following statements is supported *by the data given in the graph?*

A. In some countries, people brush their teeth more frequently than in other countries.

B. The more sugar people eat, the more likely they are to get caries.

C. In recent years, the rate of caries has increased in many countries.

D. In recent years, the consumption of sugar has increased in many countries.

### Scoring and Comments on Question 2.2

#### Full Credit

Code 1: B. The more sugar people eat, the more likely they are to get caries.

### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Health

**Setting:** Personal

This item performed reasonably well in the field trial. It was of medium difficulty with an adequate discrimination.

Like the previous item, the knowledge category is "Scientific explanations" and the competency involved is "Using scientific evidence". This time, however, the data (evidence) is provided in a graphical form. To interpret the graph correctly, a clear understanding of what variables are plotted is required.

---

## Question 2.3

A country has a high number of decayed teeth per person.

Can the following questions about tooth decay in that country be answered by scientific experiments? *Circle "Yes" or "No" for each question.*

| Can this question about tooth decay be answered by scientific experiments? | Yes or No? |
|---|---|
| What would be the effect on tooth decay of putting fluoride in the water supply? | Yes / No |
| How much should a visit to the dentist cost? | Yes / No |

### Scoring and Comments on Question 2.3

### Full Credit

Code 1: Both correct: Yes, No in that order

### No Credit

Code 0: Other responses

Code 9: Missing

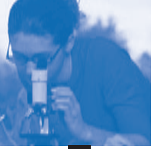
**Item type:** Complex multiple choice

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Health

**Setting:** Social





This item requires students to distinguish between questions that can be answered by scientific experimentation and those that cannot. It mainly involves applying knowledge about the methodology of science and so the knowledge category is "Scientific enquiry". The competency classification is clearly "Identifying scientific issues".

Field trial analysis placed this item in the medium difficulty range. It showed good discrimination.

### Question 2.4

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing what tooth decay bacteria look like under a microscope | ☐1 | ☐2 | ☐3 | ☐4 |
| b) | Learning about the development of a vaccine to prevent tooth decay | ☐1 | ☐2 | ☐3 | ☐4 |
| c) | Understanding how sugar-free foods can cause tooth decay | ☐1 | ☐2 | ☐3 | ☐4 |

**Item type:** Attitudinal

**Attitude:** Interest

This item was designed to assess a student's interest in scientific aspects of tooth decay. Like all attitudinal items, it comes at the end of the unit so that the students are familiar with the context prior to being asked their opinions.

Exploratory factor analysis of the field trial results showed all three statements loading significantly on an "interest" dimension.



# SCIENCE UNIT 3

## Hot Work

### Question 3.1

Peter is working on repairs to an old house. He has left a bottle of water, some metal nails, and a piece of timber inside the boot of his car. After the car has been out in the sun for three hours, the temperature inside the car reaches about 40 °C.

What happens to the objects in the car? *Circle "Yes" or "No" for each statement.*

| Does this happen to the object(s)? | Yes or No? |
|---|---|
| They all have the same temperature. | Yes / No |
| After some time the water begins to boil. | Yes / No |
| After some time the metal nails begin to glow red. | Yes / No |

### Scoring and Comments on Question 3.1

**Full Credit**

Code 1: All three correct: Yes, No, No, in that order

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

This unit has a different "style" to other units in this collection of examples. It has no common stimulus and was one of a number of units designed to promote knowledge about major misconceptions of fundamental concepts. Only one such unit was included in the main study due to space constraints and the lack of information that many of the items provided about students' overall levels of scientific literacy.

The field trial results for this item showed significant misconceptions among students with less then 20% correctly answering that "They all have the same temperature". There was virtually no difference in mean ability between those answering this option correctly and those who did not answer it correctly. Males were more likely to answer it correctly than were females. The second and third statements were each answered correctly by about 75% of students.



## Question 3.2

For drinks during the day, Peter has a cup of hot coffee, at a temperature of about 90 °C, and a cup of cold mineral water, with a temperature of about 5 °C. The cups are of identical type and size and the volume of each drink is the same. Peter leaves the cups sitting in a room where the temperature is about 20 °C.

What are the temperatures of the *coffee* and the *mineral water* likely to be after 10 minutes?

A. 70 °C and 10 °C

B. 90 °C and 5 °C

C. 70 °C and 25 °C

D. 20 °C and 20 °C

### Scoring and Comments on Question 3.2

#### Full Credit

Code 1: A. 70 °C and 10 °C

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

Like a many of the "misconception" items, this item has a rather contrived context and this was an additional reason for their absence from the main study.

In the field trial the item showed adequate discrimination and about 50% of students answered it correctly.



### Question 3.3

How much interest do you have in the following information?

*Tick only one box in each row.*

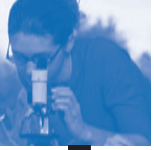
| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Understanding how the shape of the cup influences the rate at which coffee cools | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| b) | Learning about the different arrangements of atoms in wood, water and steel | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| c) | Knowing why different solids conduct heat at different rates | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |

**Item type:** Attitudinal

**Attitude:** Interest in learning about science

In this item, interest in science is shown by students expressing the degree to which they want to acquire information relating to how the structure of materials affects the transfer of heat. The item comes at the end of the unit to allow students to become familiar with the context prior to being asked for their opinions.

Exploratory factor analysis showed all three statements loading significantly on an "interest" dimension. In comparison with other contexts used in the field trial, interest shown in acquiring scientific information relevant to this context was very low. This was particularly so for the first two statements.





# SCIENCE UNIT 4

## Mousepox

There are many types of pox viruses that cause pox diseases in animals. Each type of virus usually infects only one animal species. A magazine has reported that a scientist has used genetic engineering to modify the DNA of mousepox. The altered virus kills all the mice it infects.

The scientist says research on modifying viruses is necessary in order to control pests that damage human food. Critics of the research say viruses could escape from laboratories and infect other animals. They are also worried that a modified pox virus for one species could infect other species, especially humans.

Humans are infected with a pox virus called smallpox. Smallpox kills most people it infects. While it is thought that this disease has been eliminated from the general population, smallpox virus samples are kept in laboratories around the world.

---

### Question 4.1

Critics have expressed concern that the mousepox virus could infect species other than mice. Which one of the following reasons is the *best* explanation for this concern?

A. The genes of smallpox virus and the genes of modified mousepox virus are identical.

B. A mutation in mousepox DNA might allow the virus to infect other animals.

C. A mutation could make the mousepox DNA identical to smallpox DNA.

D. The number of genes in mousepox virus is the same as in other pox viruses.

### Scoring and Comments on Question 4.1

**Full Credit**

Code 1: B. A mutation in mousepox DNA might allow the virus to infect other animals.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Global

The context of genetic modification and mutation, involving the effects of the introduction of new species into established ecosystems, and the dangers of disease “jumping” across species, is highly topical and important.

However, this item did not work in the field trial. Its discrimination was poor, and variability between countries was at an unacceptable level – indicating, perhaps, that the science involved is not part of the curricula in some countries. In addition, the mean ability of students who chose distractor C was close to the mean ability of the students who chose the correct option (B). Consequently, this item was not considered for inclusion in the main study.

---





### Question 4.2

A person who criticised the research was worried that the modified mousepox virus might escape from a laboratory. This virus could cause the extinction of some species of mice.

Are the following outcomes likely if some species of mice become extinct?

*Circle "Yes" or "No" in each case.*

| Is this outcome likely if some species of mice become extinct? | Yes or No? |
|---|---|
| Some food chains could be affected. | Yes / No |
| Domestic cats could die for lack of food. | Yes / No |
| Plants whose seeds are eaten by mice could temporarily increase in number. | Yes / No |

### Scoring and Comments on Question 4.2

#### Full Credit

Code 1: All three correct: Yes, No, Yes in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Global

Knowledge of food webs is drawn upon in answering this question correctly. The student is required to predict effects of removing components of a food web rather than use evidence in drawing a conclusion. Thus the classification is "Explaining phenomena scientifically – Knowledge of living systems".

In the field trial, the item exhibited adequate discrimination and was of medium difficulty.

---

### Question 4.3

One company is trying to develop a virus that will make mice sterile (i.e. unable to have babies). Such a virus could help control the number of mice.

Suppose the company is successful. Should the following questions be answered by research before releasing the virus?

*Circle "Yes" or "No" in each case.*

| Should this question be answered before releasing the virus? | Yes or No? |
|---|---|
| What is the best method for spreading the virus? | Yes / No |
| How soon will mice develop immunity to the virus? | Yes / No |
| Will the virus affect other animal species? | Yes / No |



### Scoring and Comments on Question 4.3

#### Full Credit

Code 1: All three correct: Yes, Yes, Yes

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically / Identifying scientific issues

**Knowledge category:** Living systems (Knowledge of science) / Scientific enquiry (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Social

This item performed quite well in the field trial with an adequate discrimination. Item difficulty was towards the lower end of the medium range.

The item was not considered for inclusion in the main study, however, because it assesses both "Knowledge about science" and "Knowledge of science" in significant amounts. Knowledge of living systems is needed to decide whether the questions should be answered before the virus is released; knowledge about science methodology is needed to decide whether the questions could be answered by scientific research.

---

### Question 4.4

How much interest do you have in the following information?

*Tick only one box in each row.*

| | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|
| a) Learning about the structure of viruses | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| b) Knowing how viruses mutate (change) | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| c) Understanding better how the body defends itself against viruses | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |

**Item type:** Attitudinal

**Attitude:** Interest

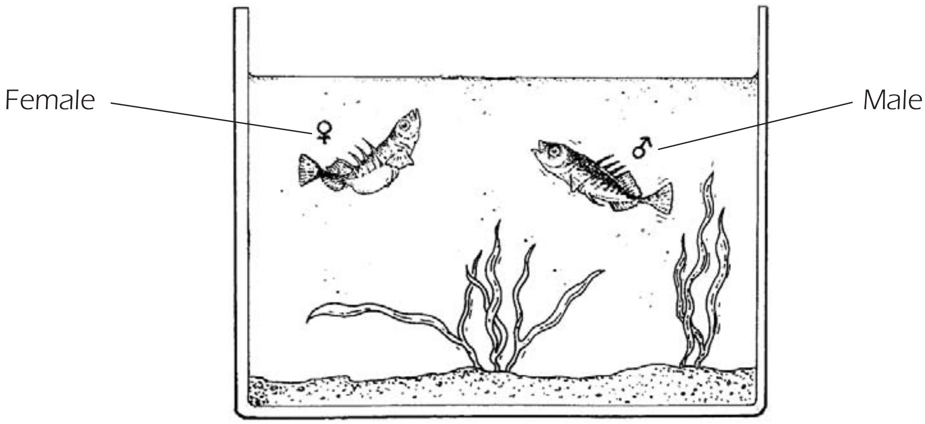
This item attempts to assess students' interest in elements of the context concerned with the behaviour of viruses and how bodily defences operate against viruses. Like all attitudinal items, it comes at the end of the unit to allow students to become familiar with the context prior to being asked for their opinions.

Exploratory factor analysis of the field trial results showed all three statements loaded significantly on an "interest" dimension. More interest was shown in understanding better how the body defends itself against viruses (statement c) than in the other two tasks.



## Stickleback Behaviour

The stickleback is a fish that is easy to keep in an aquarium.

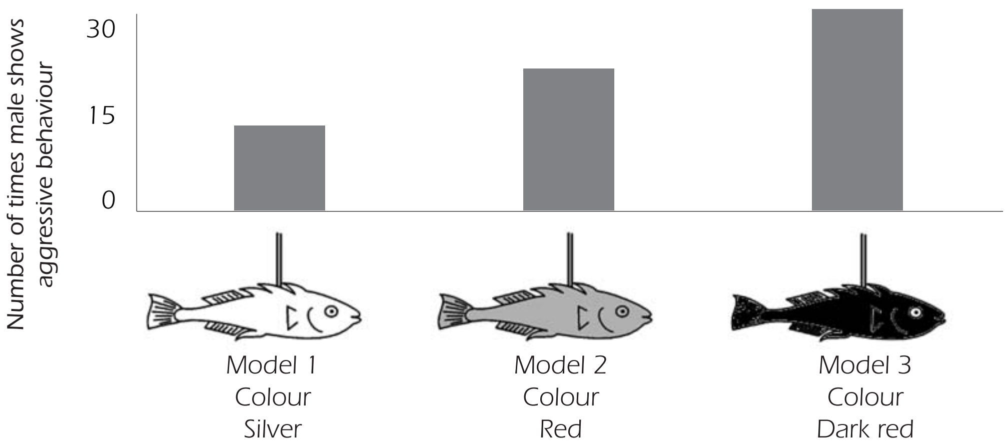



- During the breeding season the male stickleback's belly turns from silver-coloured to red.
- The male stickleback will attack any competing male that comes into his territory, and try to chase it away.
- If a silver-coloured female approaches, he will try to guide her to his nest so she will lay her eggs there.

In an experiment a student wants to investigate what will make the male stickleback show aggressive behaviour.

A male stickleback is alone in the student's aquarium. The student has made three wax models attached to pieces of wire. He hangs them separately in the aquarium for the same amount of time. Then the student counts the number of times the male stickleback reacts aggressively by pushing against the wax figure.

The results of this experiment are shown below.





## Question 5.1

What is the question that this experiment is attempting to answer?

### Scoring and Comments on Question 5.1

#### Full Credit

Code 1: What colour elicits the strongest aggressive behaviour by the male stickleback?

- Does the male stickleback react more aggressively to a red-coloured model than to a silver-coloured one?
- Is there a relationship between colour and aggressive behaviour?
- Does the colour of the fish cause the male to be aggressive?
- What fish colour does the stickleback find most threatening?

#### No Credit

Code 0: Other responses (including all responses that do not refer to the colour of the stimulus/model/fish).

- What colour will elicit aggressive behaviour in the male stickleback. [No comparative aspect.]
- Does the colour of the female stickleback determine the aggressiveness of the male? [The first experiment is not concerned with the gender of the fish.]
- Which model does the male stickleback react to most aggressively? [Specific reference must be made to the colour of the fish/model.]

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

All relevant information about the experiment is supplied and hence the "Knowledge about science" classification. The context classification ("Personal"; "Frontiers of science and technology") is in accord with the framework descriptor "expanding one's understanding of the natural world".

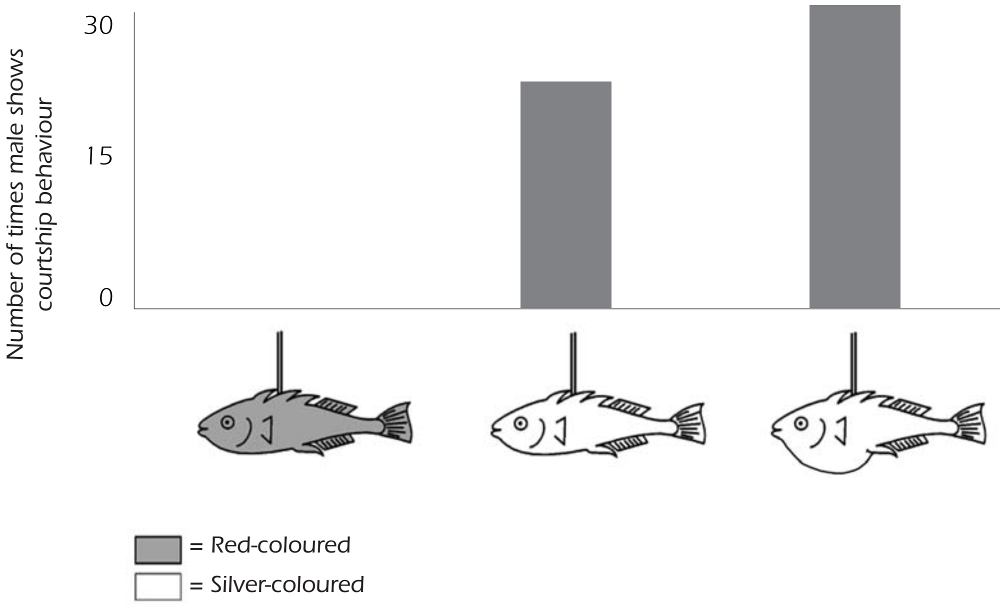
In the field trial the item demonstrated an adequate discrimination but it was generally difficult with about 25% of students gaining credit. This unit was not included in the main study because it was considered less relevant to 15-year-olds' daily lives than other units vying for inclusion, and because of its high overall reading load.

### Question 5.2

During breeding time, if the male stickleback sees a female he will try to attract the female with courtship behaviour that looks like a little dance. In a second experiment, this courtship behaviour is investigated.

Again, three wax models on a piece of wire are used. One is red-coloured; two are silver-coloured with one having a flat belly and the other a round belly. The student counts the number of times (in a given amount of time) that the male stickleback reacts to each model by showing courtship behaviour.

The results of this experiment are shown below.




Three students each draw a conclusion based on the results of this second experiment.

Are their conclusions correct according to the information given in the graph?

*Circle "Yes" or "No" for each conclusion.*

| Is this conclusion correct according to the information in the graph? | Yes or No? |
|---|---|
| The red colour causes courtship behaviour by the male stickleback. | Yes / No |
| A flat-bellied female stickleback causes most courtship behaviour from a stickleback male. | Yes / No |
| The male stickleback shows courtship behaviour more often to a round-bellied female than to a flat-bellied female. | Yes / No |

### Scoring and Comments on Question 5.2

**Full Credit**

Code 1: All three correct: No, No, Yes in that order

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

In interpreting the data that is presented graphically, the student is explaining the meanings inherent in that data and does not have to draw on any external information. Consequently, this item is classified as "Knowledge about science", category: "Scientific explanations".

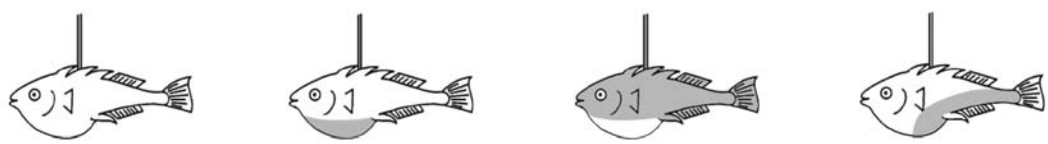
In the field trial, the item was found to be relatively easy with very good discrimination. Females were more likely to answer it correctly than were males.

---

## Question 5.3

Experiments have shown that male sticklebacks react with aggressive behaviour to models with a red *belly*, and with courtship behaviour to models with a silver *belly*.

In a third experiment, the following four models were used in turn:

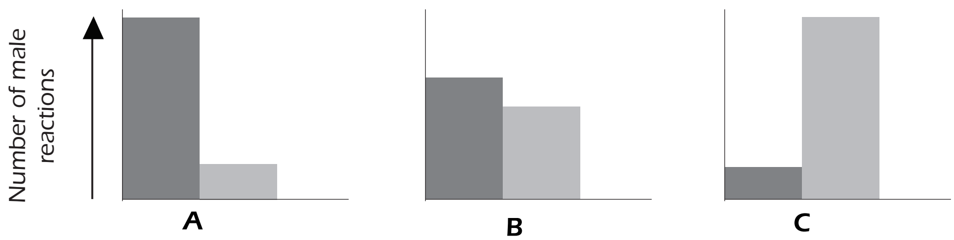


= Red-coloured

= Silver-coloured

The three diagrams below show possible reactions of a male stickleback to each of the above models.

Which one of these reactions would you predict for each of the four models?




= Number of aggressive behaviours

= Number of courtship behaviours



Fill in either A, B or C as the result for each model.

| | Reaction |
|---|---|
| Model 1 | |
| Model 2 | |
| Model 3 | |
| Model 4 | |

### Scoring and Comments on Question 5.3

#### Full Credit

Code 2: All four correct: C, A, C, B in that order

#### Partial Credit

Code 1: Three of the four entries correct

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Closed-constructed response

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

***Application area:*** Frontiers of science and technology

***Setting:*** Personal

This item did not perform well in the field trial. Discrimination was low and only about one-third of all students obtained full or partial credit. Unfortunately, data was not available on which part(s) caused most difficulty. Females were more likely to answer the item correctly than were males.



# SCIENCE UNIT 6

## Tobacco Smoking

Tobacco is smoked in cigarettes, cigars and pipes. Research shows that tobacco-related diseases kill nearly 13 500 people worldwide every day. It is predicted that, by 2020, tobacco-related diseases will cause 12% of all deaths globally.

Tobacco smoke contains many harmful substances. The most damaging substances are tar, nicotine and carbon monoxide.

---

### Question 6.1

Tobacco smoke is inhaled into the lungs. Tar from the smoke is deposited in the lungs and this prevents the lungs from working properly.

Which one of the following is a function of the lungs?

A. To pump oxygenated blood to all parts of your body

B. To transfer some of the oxygen that you breathe to your blood

C. To purify your blood by reducing the carbon dioxide content to zero

D. To convert carbon dioxide molecules into oxygen molecules

### Scoring and Comments on Question 6.1

**Full Credit**

Code 1: B. To transfer oxygen from the air that you breathe to your blood

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal

This unit was rated highly for inclusion in the main study by country representatives. This particular item does not deal directly with tobacco smoking but with the functioning of the lungs. Students were required to correctly draw on knowledge of lung function in order to gain credit so the classification is "Knowledge of science", category: "Living systems".

The field trial showed the item to be relatively easy with adequate discrimination. Males were somewhat more likely to answer it correctly than were females.

---



### Question 6.2

Tobacco smoking increases the risk of getting lung cancer and some other diseases.

Is the risk of getting the following diseases increased by tobacco smoking?

*Circle "Yes" or "No" in each case.*

| Is the risk of contracting this disease increased by smoking? | Yes or No? |
|---|---|
| Bronchitis | Yes / No |
| HIV/AIDS | Yes / No |
| Chicken pox | Yes / No |

### Scoring and Comments on Question 6.2

#### Full Credit

Code 1: All three correct: Yes, No, No in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal

This item meets the requirement of relevance for 15-year-olds. Students need to know that HIV/AIDS and chicken pox are caused by viruses, whereas bronchitis is a lung affliction and so is likely to be more common in smokers than non-smokers.

The item did not perform satisfactorily in the field trial. While being generally easy with around 70% of students responding correctly, there was, nevertheless, a considerable amount of variation in difficulty between countries. For a number of countries discrimination was very low, while females were more likely to answer the item correctly than were males.



### Question 6.3

Some people use nicotine patches to help them to give up smoking. The patches are put on skin and release nicotine into the blood. This helps to relieve cravings and withdrawal symptoms when people have stopped smoking.

To study the effectiveness of nicotine patches, a group of 100 smokers who want to give up smoking is chosen randomly. The group is to be studied for six months. The effectiveness of the nicotine patches is to be measured by finding out how many people in the group have not resumed smoking by the end of the study.

Which one of the following is the *best* experimental design?

A. All the people in the group wear the patches.

B. All wear patches except one person who tries to give up smoking without them.

C. People choose whether or not they will use patches to help give up smoking.

D. Half are randomly chosen to use patches and the other half do not use them.

### Scoring and Comments on Question 6.3

#### Full Credit

Code 1: D. Half are randomly chosen to use patches and the other half do not use them.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Health

**Setting:** Personal / Social

It is arguable whether this item is concerned with maintenance of personal health (and so the setting classification would be "Personal") or community health (setting: "Social").

A good understanding of a matched control group in experimental design is required for a correct response to this item. The field trial showed the item to be of medium difficulty but with good discrimination. Distractor B was found to be somewhat weaker than the other two distractors. Females were more likely to answer the item correctly than were males.



### Question 6.4

Various methods are used to influence people to stop smoking.

Are the following ways of dealing with the problem based on *technology*?

*Circle "Yes" or "No" in each case.*

| Is this method of reducing smoking based on technology? | Yes or No? |
|---|---|
| Increase the cost of cigarettes. | Yes / No |
| Produce nicotine patches to help make people give up cigarettes. | Yes / No |
| Ban smoking in public areas. | Yes / No |

### Scoring and Comments on Question 6.4

#### Full Credit

Code 1: All three correct: No, Yes, No in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Identifying scientific issues

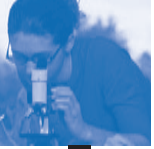
**Knowledge category:** Technology systems (Knowledge of science)

**Application area:** Health

**Setting:** Social

This item was designed to assess students' understanding of what is involved in technology. The options describe an economic, a technological/chemical, and a legislative approach to deterring people from smoking. Knowledge of the role of science-based technology is classified in the framework as "Knowledge about science", category: "Technology systems"

The item performed well in the field trial. The analysis showed it to be of medium difficulty with good discrimination.





### Question 6.5

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing how tar in tobacco reduces lung efficiency | ☐1 | ☐2 | ☐3 | ☐4 |
| b) | Understanding why nicotine is addictive | ☐1 | ☐2 | ☐3 | ☐4 |
| c) | Learning how the body recovers after stopping smoking | ☐1 | ☐2 | ☐3 | ☐4 |

**Item type:** Attitudinal

**Attitude:** Interest

Exploratory factor analysis showed a general loading on an "interest in science" dimension for the first statement, but there was a greater loading on a dimension thought to represent interest in/concern for health and safety. The other two statements performed inconsistently across countries. Even if the unit had been included in the main study, this item would have been omitted on these grounds.



## SCIENCE UNIT 7

### Starlight

Toshio likes to look at stars. However, he cannot observe stars very well at night because he lives in a large city.

Last year Toshio visited the countryside where he observed a large number of stars that he cannot see when he is in the city.

---

### Question 7.1

Why can many more stars be observed in the countryside than in large cities?

A. The moon is brighter in cities and blocks out the light from many stars.

B. There is more dust to reflect light in country air than in city air.

C. The brightness of city lights makes many stars hard to see.

D. The air is warmer in cities due to heat emitted by cars, machinery and houses.

#### Scoring and Comments on Question 7.1

**Full Credit**

Code 1: C. The brightness of city lights makes many stars hard to see.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Earth and space systems (Knowledge of science)

**Application area:** Environment

**Setting:** Social

In this item, students' knowledge of the effect of extraneous light on their ability to resolve light from stars is needed to select the correct response. Consequently, the classification is "Explaining phenomena scientifically – Earth and space systems".

The item performed quite well in the field trial showing adequate discrimination and with minimal evidence of gender or cultural bias. It was answered correctly by about 65% of students.

---



## Question 7.2

Toshio uses a telescope with a large diameter lens in order to observe stars of low brightness.

Why does using a telescope with a large diameter lens make it possible to observe stars of low brightness?

A. The larger the lens the more light is collected.

B. The larger the lens the more it magnifies.

C. Larger lenses allow more of the sky to be seen.

D. Larger lenses can detect the dark colours in stars.

### Scoring and Comments on Question 7.2

#### Full Credit

Code 1: A. The larger the lens the more light is collected.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

Country representatives rated this item rather low for inclusion in the main study compared with other field trial items. The item showed adequate discrimination and was of medium difficulty in the field trial. Surprisingly, distractor D (selected by 45% of students) was far more popular than the correct answer A (selected by about 30% of students). Males were more likely to answer the item correctly than were females.



## SCIENCE UNIT 8

### Ultrasound

In many countries, images can be taken of a foetus (developing baby) by ultrasound imaging (echography). Ultrasounds are considered safe for both the mother and the foetus.

The doctor holds a probe and moves it across the mother's abdomen. Ultrasound waves are transmitted

into the abdomen. Inside the abdomen they are reflected from the surface of the foetus. These reflected waves are picked up again by the probe and relayed to a machine that can produce an image.

---

### Question 8.1

To form an image the ultrasound machine needs to calculate the *distance* between the foetus and the probe.

The ultrasound waves move through the abdomen at a speed of 1540 m/s.
What measurement must the machine make so that it can calculate the distance?

#### Scoring and Comments on Question 8.1

**Full Credit**

Code 1: It must measure the time the ultrasound wave takes to travel from the probe to the surface of the foetus and reflect back.

- The time of travel of the wave.
- The time.
- Time. Distance = speed / time. [Although the formula is incorrect, the student has correctly identified "time" as the missing variable.]
- It must find when the ultrasound finds the baby.

**No Credit**

Code 0: Other responses

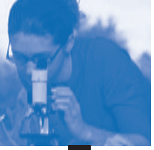
- The distance.

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)





**Application area:** Frontiers of science and technology

**Setting:** Personal

The context of this unit has strong relevance to the lives of citizens – in particular, females. The first item differs from the remaining two items in the unit, however, because it is concerned with the science behind the technology rather than the effect or function of technology. Males were more likely to answer it correctly than were females.

The item was answered correctly by only about 20% of students in the field trial, but displayed very good discrimination.

---

## Question 8.2

An image of a foetus can also be obtained using X-rays. However, women are advised to avoid having X-rays of their abdomens during pregnancy.

Why should a woman avoid having her abdomen X-rayed during pregnancy *in particular*?

### Scoring and Comments on Question 8.2

#### Full Credit

Code 1: X-rays are harmful to the foetus.

- X-rays hurt the foetus.
- X-rays might cause a mutation in the foetus.
- X-rays can cause birth defects in the foetus.
- Because the baby could get some radiation.

#### No Credit

Code 0: Other responses

- X-rays do not give a clear picture of the foetus.
- X-rays emit radiation.
- The child can get Down syndrome.
- Radiation is harmful. [This is not enough. Potential harm to the foetus (baby) must be explicitly mentioned.]
- They may make it harder for her to have another baby. [This is a reason for avoiding over-exposure to X-rays in general.]

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal

Field trial results showed this item to be easy with around 75% of students answering it correctly. Discrimination was adequate and there were no obvious cross-country issues. Not surprisingly, females were more likely to answer the item correctly than were males.

---



### Question 8.3

Can ultrasound examinations of expectant mothers provide answers to the following questions?

*Circle "Yes" or "No" for each question.*

| Can an ultrasound examination answer this question? | Yes or No? |
|---|---|
| Is there more than one baby? | Yes / No |
| What colour are the baby's eyes? | Yes / No |
| Is the baby about the right size? | Yes / No |

### Scoring and Comments on Question 8.3

#### Full Credit

Code 1: All three correct: Yes, No, Yes in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal

This item was answered correctly by about 70% of students and so can be considered to be quite easy. Females were much more likely to answer it correctly than were males.

The item can be answered by drawing on knowledge of the nature of ultrasound and what it can detect – hence the classification "Physical systems". However, the item can also be answered from personal familiarity with the technology of ultrasound imaging, making it easier for such students. This contributed to the decision not to include the unit in the main study.

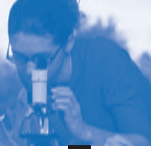
---

### Question 8.4

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Understanding how ultrasound can penetrate your body without harming it | ☐1 | ☐2 | ☐3 | ☐4 |
| b) | Learning more about the differences between X-rays and ultrasound | ☐1 | ☐2 | ☐3 | ☐4 |
| c) | Knowing about other medical uses of ultrasound | ☐1 | ☐2 | ☐3 | ☐4 |





**Item type:** *Attitudinal*

***Attitude:*** *Interest in learning about science*

*Exploratory factor analysis showed all three statements loading on an "interest in science" dimension, but each statement loaded more strongly on a dimension presumed to measure interest in/concern for health and safety. That this was true for statement b) as well as for statements a) and c), even though it is not as directly concerned with health and safety issues, is probably a consequence of the item being embedded in a context that is strongly health oriented.*



# SCIENCE UNIT 9

## Lip Gloss

The table below contains two different recipes for cosmetics you can make yourself.

The lipstick is harder than the lip gloss, which is soft and creamy.

| **Lip gloss** | **Lipstick** |
| --- | --- |
| **Ingredients:**<br>5 g castor oil<br>0.2 g beeswax<br>0.2 g palm wax<br>1 teaspoon of colouring substance<br>1 drop of food flavouring | **Ingredients:**<br>5 g castor oil<br>1 g beeswax<br>1 g palm wax<br>1 teaspoon of colouring substance<br>1 drop of food flavouring |
| **Instructions:**<br>Heat the oil and the waxes in a container placed in hot water until you have an even mixture. Then add the colouring substance and the flavouring, and mix them in. | **Instructions:**<br>Heat the oil and the waxes in a container placed in hot water until you have an even mixture. Then add the colouring substance and the flavouring, and mix them in. |

### Question 9.1

In making the lip gloss and lipstick, oil and waxes are mixed together. The colouring substance and flavouring are then added.

The lipstick made from this recipe is hard and not easy to use. How would you change the proportion of ingredients to make a softer lipstick?

### Scoring and Comments on Question 9.1

#### Full Credit

Code 1: Responses indicating that you would add less wax AND/OR add more oil

- You could use a bit less beeswax and palm wax.
- Add more castor oil.
- Put in 7 g of oil.

#### No Credit

Code 0: Other responses

- Heat the mixture for longer which will soften it.
- By not heating the waxes as much. [The question asks how you would change the proportion of ingredients.]

Code 9: Missing

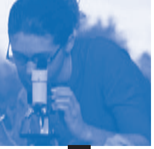
**Item type:** Open-constructed response

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal





The context of cosmetics has everyday relevance for students of this age group, although it could be expected that this unit would generate more interest among females than males.

This item can be answered by comparing the quantities of ingredients used in the two recipes to conclude why one recipe produces a softer substance than the other one. The item is therefore classified as "Knowledge about science", category: "Scientific explanations". However, it helps to have knowledge of the properties of the main ingredients (oil and wax) and a case can be made for classifying the item as "Knowledge of science", category: "Physical systems"; competency: "Explaining phenomena scientifically".

In the field trial, about 65% of students answered the item correctly and it displayed good discrimination. Females were much more likely to answer it correctly than were males.

---

## Question 9.2

Oils and waxes are substances that will mix well together. Oils cannot be mixed with water, and waxes are not soluble in water.

Which one of the following is most likely to happen if a lot of water is splashed into the lipstick mixture while it is being heated?

A. A creamier and softer mixture is produced.

B. The mixture becomes firmer.

C. The mixture is hardly changed at all.

D. Fatty lumps of the mixture float on the water.

### Scoring and Comments on Question 9.2

#### Full Credit

Code 1: D. Fatty lumps of the mixture float on the water.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

This item has less everyday relevance than other items in this unit. Students must reason from the information provided in the stimulus in selecting an appropriate prediction from those on offer. Thus the item is classified as "Knowledge about science", competency: "Scientific explanations".

About 70% of students answered the item correctly. As for Question 9.1, females were much more likely to answer it correctly than were males.

---



### Question 9.3

When substances called emulsifiers are added, they allow oils and waxes to mix well with water.

Why does soap and water remove lipstick?

A. Water contains an emulsifier that allows the soap and lipstick to mix.

B. The soap acts as an emulsifier and allows the water and lipstick to mix.

C. Emulsifiers in the lipstick allow the soap and water to mix.

D. The soap and lipstick combine to form an emulsifier that mixes with the water.

### *Scoring and Comments on Question 9.3*

**Full Credit**

Code 1: B. The soap acts as an emulsifier and allows the water and lipstick to mix.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

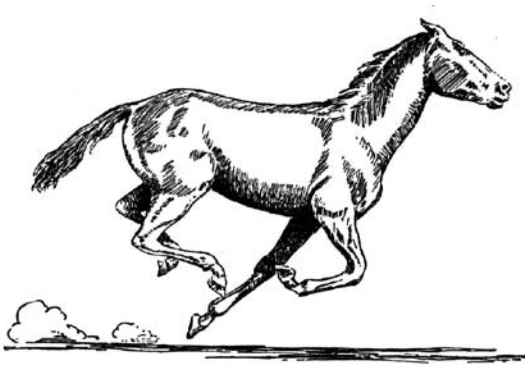
Unlike other items in the unit, there was no discernible difference in performance between males and females on this item in the field trial. Like the previous item, an explanation that accords with the information supplied has to be selected from the four options. Consequently, this item has the same knowledge and competency classifications.

The item performed well in the field trial with good discrimination and had medium difficulty.

# SCIENCE UNIT 10

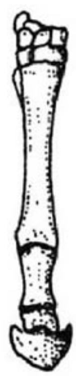
## Evolution

Most horses today are streamlined and can run fast.



Scientists have found the fossil skeletons of animals that are similar to horses. They consider them to be the ancestors of the present-day horse. The scientists have also been able to determine the period during which the fossil species were living.

The table below provides information on three of these fossils and on the present-day horse.

| ANIMAL NAME: | HYRACOTHERIUM | MESOHIPPUS | MERYCHIPPUS | EQUUS (present-day horse) |
|---|---|---|---|---|
| Period of existence: | 55 to 50 million years ago | 39 to 31 million years ago | 19 to 11 million years ago | 2 million years ago to the present day |
| Skeleton of the leg (same scale): |  |  |  |  |

### Question 10.1

What information *in the table* is strong evidence that present-day horses may have evolved over time from the other three animals?

#### Scoring and Comments on Question 10.1

**Full Credit**

Code 1: Responses that refer to gradual change (progression) in leg skeleton structure over time

- The leg skeletons are much the same but have gradually changed.
- The digits/toes fused during the period 55 to 2 million years ago.
- The number of digits has decreased.



### No Credit

Code 0: Other responses

- The leg has changed. [Not specific enough.]
- They are called Hippus.
- Genetic mutations have caused the transformations. [Correct, but does not answer the question.]
- The leg bones are similar. [Need to mention or imply "gradual change".]

Code 9: Missing

**Item type:** Open-constructed response

**Competency**: Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Global

While the knowledge assessed in this unit may not rate highly on "everyday relevance", it is concerned with one of the most important "big issues" in science and the unit would have been included in the main study if the items had performed more satisfactorily in the field trial.

This item involves a comparative analysis of the data in the table with a view to providing a scientific explanation. Hence it is classified as "Knowledge about science", category: "Scientific explanations". The version included here has been revised since the field trial because of difficulties encountered in reliably coding the field trial version.

---

## Question 10.2

What further research can scientists undertake to find out how horses have evolved over time?

*Circle "Yes" or "No" for each of these statements.*

| Would this research help find out how horses have evolved over time? | Yes or No? |
|---|---|
| Compare the number of horses living at different periods. | Yes / No |
| Search for skeletons belonging to ancestors of the horse that lived 50 to 40 million years ago. | Yes / No |

### Scoring and Comments on Question 10.2

### Full Credit

Code 1: Both correct: No, Yes in that order

### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)



**Application area:** Frontiers of science and technology

**Setting:** Global

This item asks what further evidence should be sought to help answer a scientific question. To this end, some familiarity with evolution or natural selection (i.e. Knowledge of science) is required. However, on balance, the main knowledge requirement is recognising whether the research referred to is feasible. Hence the item is classified as "Knowledge about science", category: "Scientific enquiry".

The item performed quite well in the field trial with adequate discrimination and without significant cross-country or gender issues. It was of medium difficulty.

---

### Question 10.3

Which one of the following statements best applies to the scientific theory of evolution?

A. The theory cannot be believed because it is not possible to see species changing.

B. The theory of evolution is possible for animals but cannot be applied to humans.

C. Evolution is a scientific theory that is currently based on extensive evidence.

D. Evolution is a theory that has been proven to be true by scientific experiments.

### Scoring and Comments on Question 10.3

#### Full Credit

Code 1: C. Evolution is a scientific theory that is currently based on extensive evidence.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Global

Field trial results exposed problems with this item. Distractor D proved almost as popular as the correct option (C). In addition, there was considerable variability in item difficulty across countries and discrimination was very low in a number of countries. In the version included here, the wording of option C has been changed slightly from the field trial version.

---



## Question 10.4

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing how fossils can be identified | □ 1 | □ 2 | □ 3 | □ 4 |
| b) | Learning more about the development of the theory of evolution | □ 1 | □ 2 | □ 3 | □ 4 |
| c) | Understanding better the evolution of the present-day horse | □ 1 | □ 2 | □ 3 | □ 4 |

**Item type:** Attitudinal

**Attitude:** Interest in learning about science

This item was designed to assess students' interest in the science of evolution. Like all attitudinal items, it comes at the end of the unit so that students are familiar with the context prior to being asked their opinions.

Exploratory factor analysis showed significant loadings on an "interest" dimension for all three statements, with negligible loadings on the other dimensions explored. Much less interest was shown in the last statement compared with the other two statements.



# SCIENCE UNIT 11

## Bread Dough

To make bread dough, a cook mixes flour, water, salt and yeast. After mixing, the dough is placed in a container for several hours to allow the process of fermentation to take place. During fermentation, a chemical change occurs in the dough: the yeast (a single-celled fungus) helps to transform the starch and sugars in the flour into carbon dioxide and alcohol.

---

### Question 11.1

Fermentation causes the dough to rise. Why does the dough rise?

A. The dough rises because alcohol is produced and turns into a gas.

B. The dough rises because of single-celled fungi reproducing in it.

C. The dough rises because a gas, carbon dioxide, is produced.

D. The dough rises because fermentation turns water into a vapour.

### Scoring and Comments on Question 11.1

#### Full Credit

Code 1: C. The dough rises because a gas, carbon dioxide, is produced.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

The field trial showed that this item had adequate discrimination and was of medium difficulty. However, there were a number of countries in which the mean ability of students who chose option A was close to the mean ability of the students who chose the correct option (C).



### Question 11.2

A few hours after mixing the dough, the cook weighs the dough and observes that its weight has decreased.

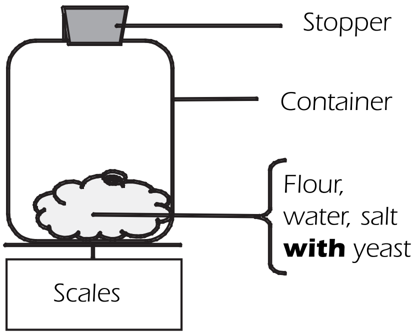
The weight of the dough is the same at the start of each of the four experiments shown below. Which *two* experiments should the cook compare to test if the *yeast* is the cause of the loss of weight?

A. The cook should compare experiments 1 and 2.

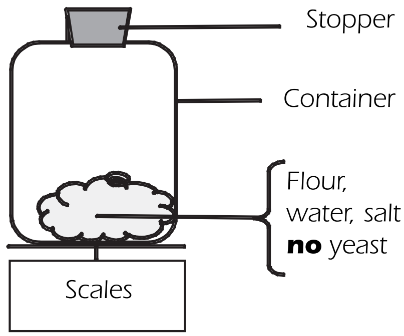
B. The cook should compare experiments 1 and 3.

C. The cook should compare experiments 2 and 4.

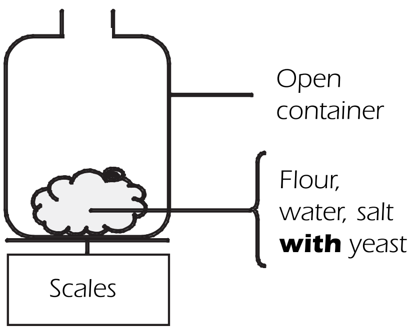
D. The cook should compare experiments 3 and 4.



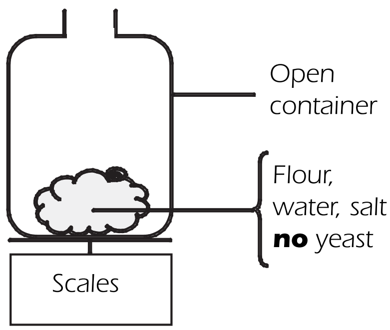

**Experiment 1**




**Experiment 2**




**Experiment 3**




**Experiment 4**

### Scoring and Comments on Question 11.2

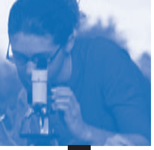
#### Full Credit

Code 1: D. The cook should compare experiments 3 and 4.

#### No Credit

Code 0: Other responses

Code 9: Missing





**Item type:** Multiple choice

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

In selecting an appropriate response, the student must identify the variable that needs to be changed (yeast presence/absence) and the variables that need controlling (the other ingredients). The student must also recognise that the presence of the stopper would prevent gas from escaping in contrast with the situation being simulated. Consequently, this item is classified as "Knowledge about science", category: "Scientific enquiry" and competency "Identifying scientific issues".

Only about one-quarter of students answered the item correctly in the field trial, and it displayed poor discrimination.

---

## Question 11.3

In the dough, yeast helps to transform starch and sugars in the flour. A chemical reaction occurs during which carbon dioxide and alcohol form.

Where do the *carbon atoms* that are present in carbon dioxide and alcohol come from? Circle "Yes" or "No" for each of the following possible explanations.

| Is this a correct explanation of where the carbon atoms come from? | Yes or No? |
|---|---|
| Some carbon atoms come from the sugars. | Yes / No |
| Some carbon atoms are part of the salt molecules. | Yes / No |
| Some carbon atoms come from the water. | Yes / No |

### Scoring and Comments on Question 11.3

#### Full Credit

Code 1: All three correct: Yes, No, No in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

One option (explanation) has been deleted from this item since the field trial. The deleted option was poorly answered and adversely affected the item's discrimination. It is anticipated that the revised item would still be somewhat difficult but would have an adequate discrimination.

Students are required to draw on knowledge of the atomic components of sugar, salt and water in order to give a correct response. This is knowledge of (physical) science.



### Question 11.4

When the risen (leavened) dough is placed in the oven to bake, pockets of gas and vapours in the dough expand.

Why do the gas and vapours expand when heated?

A. Their molecules get bigger.

B. Their molecules move faster.

C. Their molecules increase in number.

D. Their molecules collide less frequently.

### Scoring and Comments on Question 11.4

#### Full Credit

Code 1: B. Their molecules move faster.

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

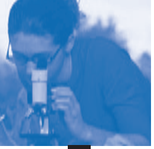
**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

This item assesses understanding of the particle model of matter. Distractors A and C represent common misconceptions and were chosen by about 25% and 20% of students respectively. About 45% of students answered the item correctly.

The item showed good discrimination overall in the field trial and was of medium difficulty. However, the variability in difficulty between countries was somewhat marked and lowered the item's priority for inclusion in the main study.



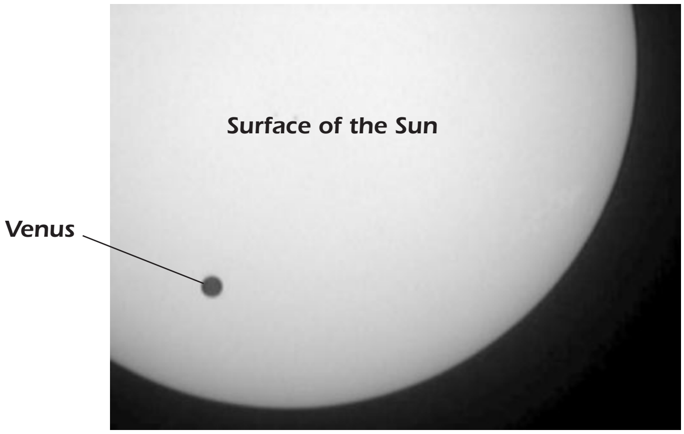


# SCIENCE UNIT 12

## Transit of Venus

On 8 June 2004, the planet Venus could be seen passing in front of the Sun when viewed from many places on Earth. This is called a "transit" of Venus and happens when its orbit takes Venus between the Sun and Earth. The previous transit of Venus occurred in 1882 and another is predicted to occur in 2012.

Below is a picture of the transit of Venus in 2004. A telescope was pointed at the Sun and the image projected onto a white card.




---

### Question 12.1

Why was the transit observed by projecting the image onto a white card, rather than by looking directly through the telescope?

A. The Sun's light was too bright for Venus to show up.

B. The Sun is big enough to see without magnification.

C. Viewing the Sun through a telescope may damage your eyes.

D. The image needed to be made smaller by projecting it onto a card.

### Scoring and Comments on Question 12.1

**Full Credit**

Code 1: C. Viewing the Sun through a telescope may damage your eyes.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal



The unit context concerns the solar system and so the knowledge area is ostensibly "Earth and space systems". However, this item involves students' awareness of the danger to the eyes of very bright sunlight and therefore it has been classified in the "Knowledge of living systems".

The unit was given a low rating for inclusion in the main study because it was felt that it would not retain its interest and relevance as time passed. In addition, as a whole it did not perform well in the field trial.

For this item, discrimination was at the low end of the acceptable range and the item was generally of medium difficulty. Males were more likely to answer it correctly than were females. In some countries, the mean ability of the students who chose distractor A was close to the mean ability of the students who chose the correct option (C).

---

### Question 12.2

When viewed from Earth, which one of the following planets can be seen in transit across the face of the Sun at certain times?

A. Mercury

B. Mars

C. Jupiter

D. Saturn

#### Scoring and Comments on Question 12.2

**Full Credit**

Code 1: A. Mercury

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Earth and space systems (Knowledge of science)

**Application area:** Frontiers of science and technology

**Setting:** Personal / Global

The setting classification of this item presents a dilemma. It is concerned with the "structure of the universe", but is hardly "frontier" knowledge in the global sense. The alternative is to regard the item as being more directed at "understanding of the natural world" and so classify it as "Personal" (along with the other items in the unit). The "Setting" and "Application area" classifications of items are often problematic and it is important to note that they will not be used for any purpose other than to ensure a variety of appropriate contexts is featured in the test.

To answer the item correctly, students need to recognise that transits can only be seen from Earth for planets between the Earth and the Sun, and know the orbital radii of Earth relative to the other planets.

In the field trial, the item proved to be difficult overall with a lot of variation in difficulty between countries. Males were more likely to answer it correctly than were females. Discrimination was at the lower end of the acceptable range.

---



## Question 12.3

Several words have been underlined in the following statement.

Astronomers predict that, as seen from Neptune, there will be a transit of Saturn across the Sun's face later this century.

Which *three* of the underlined words would be most useful in an internet or library search to find out when this transit might occur?

### Scoring and Comments on Question 12.3

**Full Credit**

Code 1: Responses referring to transit/Saturn/Neptune only
Saturn/Neptune/transit.

**No Credit**

Code 0: Other responses such as those that include four words
transit/Saturn/Sun/Neptune.
Astronomers/transit/Saturn/Neptune.

Code 9: Missing

**Item type:** Closed-constructed response

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Frontiers of science and technology

**Setting:** Personal

"Identifying keywords to search for scientific information on a given topic" is a component of the "Identifying scientific issues" competency as defined in the framework.

This item performed poorly in the field trial. It proved very difficult, with only 13% of students gaining credit. Problems arose in producing equivalent translations of the focus statement in some languages and this may have increased the item's difficulty in some countries. However, the average percentage correct in English-language countries also was 13%.





# SCIENCE UNIT 13

## Health Risk?

Imagine that you live near a large chemical factory that produces fertilisers for use in agriculture. In recent years there have been several cases of people in the area suffering from long-term breathing problems. Many local people believe that these symptoms are caused by the emission of toxic fumes from the nearby chemical fertiliser factory.

A public meeting was held to discuss the potential dangers of the chemical factory to the health of local residents. Scientists made the following statements at the meeting.

| |
|---|
| **Statement by scientists working for the chemical company**<br>"We have made a study of the toxicity of soil in the local area. We have found no evidence of toxic chemicals in the samples we have taken." |
| **Statement by scientists working for concerned citizens in the local community**<br>"We have looked at the number of cases of long-term breathing problems in the local area and compared this with the number of cases in an area far away from the chemical factory. There are more incidents in the area close to the chemical factory." |

---

### Question 13.1

The owner of the chemical factory used the statement of the scientists working for the company to argue that "the emission fumes from the factory are not a health risk to local residents".

Give one reason, other than the statement by scientists working for the concerned citizens, for *doubting* that the statement by scientists working for the company supports the owner's argument.

#### Scoring and Comments on Question 13.1

**Full Credit**

Code 1: An appropriate reason is given for doubting that the statement supports the owner's argument.

- The substance causing the breathing problems may not have been recognised as toxic.
- Breathing problems may have been caused only when chemicals were in the air, not in the soil.
- Toxic substances may change/break down with time and show up as non-toxic substances in soil.
- We do not know if the samples are representative of the area.
- Because the scientists are being paid by the company.
- The scientists feared losing their jobs.

**No Credit**

Code 0: Other responses

Code 9: Missing

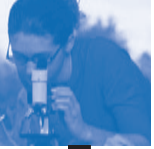
**Item type:** Open-constructed response

**Competency:** Using scientific evidence

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Hazards

**Setting:** Social





The context of this unit is particularly relevant to everyday life in that it deals with an individual's capacity to make informed judgements about a community concern by evaluating scientific information. The competency involved in both items in the unit is "Using scientific evidence".

For this first item, the science involves making judgements about relevance, adequacy of the information provided, and the credibility of that information, so the knowledge classification is "Knowledge about science", category: "Scientific enquiry".

Statistics across all countries indicated that the item performed well overall. It had good discrimination, no gender bias, and was of medium difficulty. However, there was considerable variation in difficulty range between countries. This poor country interaction, together with some discomfort amongst experts and country representatives about having to accept answers that questioned the integrity of scientific research, led to the item and unit not being considered for the main study.

---

### Question 13.2

The scientists working for the concerned citizens compared the number of people with long-term breathing problems close to the chemical factory with those in an area far away from the factory.

Describe one possible difference in the two areas that would make you think that the comparison was not a valid one.

#### Scoring and Comments on Question 13.2

**Full Credit**

Code 1: Responses should focus on possible relevant differences between the areas investigated.

- The number of people in the two areas might be different.
- One area could have better medical services than the other.
- There could be different proportions of elderly people in each area.
- There might be other air pollutants in the other area.

**No Credit**

Code 0: Other responses

- The differences between the areas might be big.

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Identifying scientific issues

**Knowledge category:** Scientific enquiry (Knowledge about science)

**Application area:** Hazards

**Setting:** Social

This item requires the student to identify variables that have not been controlled and that might influence the measured outcomes. Since the main issue is about experimental design the classification is once again "Knowledge about science", category: "Scientific enquiry".

The field trial showed the item to have good discrimination, but it was a very difficult item with only about 25% of students gaining credit.

---



### Question 13.3

How much interest do you have in the following information?

*Tick only one box in each row.*

| | | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|---|
| a) | Knowing more about the chemical composition of agricultural fertilisers | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| b) | Understanding what happens to toxic fumes emitted into the atmosphere | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |
| c) | Learning about respiratory diseases that can be caused by chemical emissions | ☐ 1 | ☐ 2 | ☐ 3 | ☐ 4 |

**Item type:** Attitudinal

**Attitude:** Interest in learning about science

Exploratory factor analysis of the field trial results showed all three statements loading significantly on an "interest in science" dimension. Despite an apparent relation to health and safety embodied in the second and third statements, there was little evidence of them loading on a dimension thought to represent interest in/concern for health and safety. A moderate amount of interest was shown in these two statements, but very little interest was shown in the first statement.



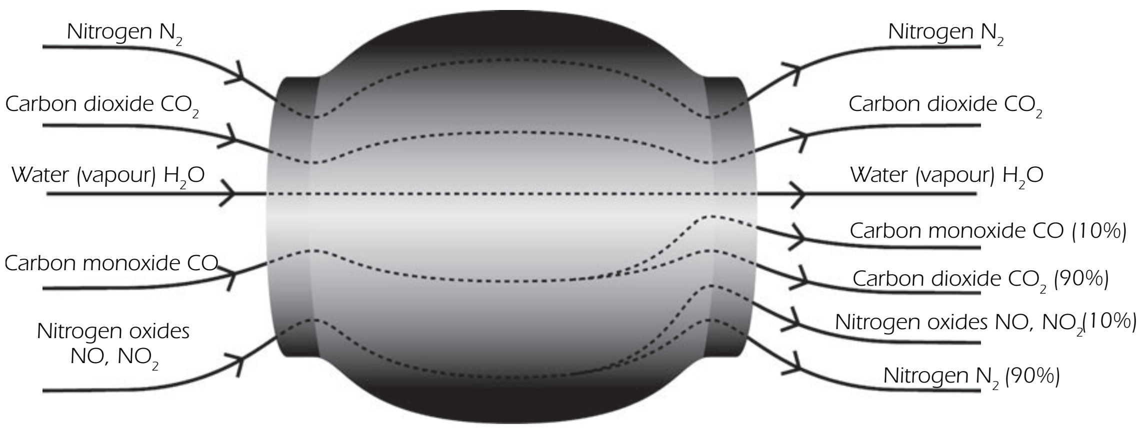
# SCIENCE UNIT 14

## Catalytic converter

Most modern cars are fitted with a catalytic converter that makes the exhaust fumes of the car less harmful to people and to the environment.

About 90% of harmful gases are converted into less harmful ones. Here are some of the gases that go into the converter and how they come out of it.

**Gases going in** **Gases coming out**




**Catalytic converter**

### Question 14.1

Use the information in the diagram above to give *an example* of how the catalytic converter makes exhaust fumes less harmful.

### Scoring and Comments on Question 14.1

#### Full Credit

Code 1: The conversion of carbon monoxide, or nitrogen oxides, to other compounds is mentioned.

- Carbon monoxide is changed into carbon dioxide.
- Nitrogen oxides are changed into nitrogen.
- It changes harmful fumes into non-harmful fumes, e.g. CO into $CO_2$ (90%).
- Carbon dioxide and nitrogen are not as harmful as carbon monoxide and nitrogen oxides.

#### No Credit

Code 0: Other responses

- The gases become less harmful.
- It purifies the carbon monoxide and nitrogen oxides. [Not specific enough.]

Code 9: Missing



**Item type:** Open-constructed response

**Competency:** Using scientific evidence

**Knowledge category:** Physical systems (Knowledge of science) / Scientific explanations (Knowledge about science)

**Application area:** Environment

**Setting:** Social

The context of atmospheric pollution by vehicle exhausts is a relevant issue of everyday living for most citizens, although that relevance is unequally shared between urban and non-urban dwellers. It was anticipated that analysis of the field trial results would show some gender difference in favour of males on the items in this unit, but this was not the case.

Students would generally bring to the unit some knowledge of what exhaust gases are toxic or environmentally harmful and which ones are not, and so the item can be been classified as “Knowledge of science”, category: “Physical systems”. However, the information supplied in the diagram indicates that carbon monoxide and nitrogen oxides are the only gases that are reduced in quantity by the catalytic converter, so it can be deduced that one or the other or both must be the harmful gases. For those making this deduction, the science in their deliberations lies in interpreting a diagram dealing with a “scientific” context. Hence, an argument can be made for a classification of “Knowledge about science”, category: “Scientific explanations” for the first item.

---

## Question 14.2

Changes take place to gases inside the catalytic converter. Explain what is happening in terms of *atoms* AND *molecules*.

### Scoring and Comments on Question 14.2

#### Full Credit

Code 2: Expresses the essential idea that atoms are rearranged to form different molecules, using **both** of these words

- Molecules break up and atoms are re-combined to form different molecules.
- Atoms rearrange to make different molecules.

#### Partial Credit

Code 1: Expresses the essential idea of rearrangement, but does not refer to both atoms and molecules OR does not distinguish sufficiently between the roles of atoms and molecules

- Atoms rearrange to make different substances.
- Molecules are changing into other molecules.
- Atoms and molecules are combining and separating to make less harmful gases. [The different roles of atoms and molecules are not sufficiently distinguished.]
- $2(NO_2) = N_2 + 2O_2$.

#### No Credit

Code 0: Other responses, including those that state no more than is given in the stimulus

- Carbon dioxide is changed into carbon monoxide.
- The molecules are being broken down into smaller atoms. [No indication that atoms are rearranged.]

Code 9: Missing



**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Environment

**Setting:** Social

This item directly assesses students understanding of what happens in a chemical reaction and the nature of the units (atoms and molecules) involved. Thus it is classified as "Knowledge of science", category: "Physical systems".

Students found it very difficult to obtain credit for the item in the field trial. About 15% obtained full credit and a similar percentage obtained partial credit.

---

## Question 14.3

Examine the gases emitted by the catalytic converter. What is one problem that engineers and scientists working on the catalytic converter should try to solve to produce less harmful exhaust fumes?

### Scoring and Comments on Question 14.3

#### Full Credit

Code 1: Acceptable responses should relate to achieving a reduction in harmful gases entering the atmosphere.

- Not all the carbon monoxide is converted into carbon dioxide.
- Not enough conversion of nitrogen oxides to nitrogen is taking place.
- Improve the percentage of carbon monoxide being converted to carbon dioxide and the percentage of nitrogen oxides being converted to nitrogen.
- The carbon dioxide produced should be captured and not allowed to escape into the atmosphere.

#### No Credit

Code 0: Other responses

- More complete conversion of the harmful gases to less harmful ones. [At least one of the harmful exhaust gases must be identified.]
- They need to try and have less fumes coming out.
- They should find a way to re-use harmful exhaust gases.
- They should try and make a vehicle that runs on a different liquid fuel.

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Using scientific evidence

**Knowledge category:** Physical systems (Knowledge of science)

**Application area:** Environment

**Setting:** Social

Answering the question requires similar knowledge and skills to those assessed by the first item in the unit (Question 14.1). Consequently, one of the two items would have been deleted if the unit had been included in the main study.

---

### Question 14.4

How much interest do you have in the following information?

*Tick only one box in each row.*

| | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|
| a) Knowing how car fuels differ in the amounts of toxic fumes they produce | ☐1 | ☐2 | ☐3 | ☐4 |
| b) Understanding more about what happens inside a catalytic converter | ☐1 | ☐2 | ☐3 | ☐4 |
| c) Learning about vehicles that do not emit toxic exhaust fumes | ☐1 | ☐2 | ☐3 | ☐4 |

**Item type:** Attitudinal

***Attitude:*** Interest in learning about science

Exploratory factor analysis showed a significant loading on an "interest in science" dimension for all three statements. In addition, there was some indication of loading on a dimension thought to represent interest in/concern for health and safety. Much more interest was shown in the last statement than in either of the other two statements.

# SCIENCE UNIT 15

## Major Surgery

Major surgery, performed in specially equipped operating theatres, is necessary to treat many diseases.

### Question 15.1

While undergoing major surgery, patients are anaesthetised so they don't feel any pain. The anaesthetic is often administered as a gas through a face mask that covers the nose and mouth.

Are the following human systems involved in the action of anaesthetic gases?

*Circle "Yes" or "No" for each system.*

| Is this system involved in the action of anaesthetic gases? | Yes or No? |
|---|---|
| Digestive system | Yes / No |
| Nervous system | Yes / No |
| Respiratory system | Yes / No |

### Scoring and Comments on Question 15.1

**Full Credit**

Code 1: All three correct: No, Yes, Yes in that order

**No Credit**

Code 0: Other responses

Code 9: Missing



**Item type:** Complex multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal / Social

Illness requiring surgery for some member of a family group is common over a period of time, so the context of this unit meets the "everyday relevance" criteria outlined in the framework. The setting classification is "Personal" or "Social" depending on whether it is judged from the perspective of a patient or a hospital.

The item displayed poor discrimination overall in the field trial, due principally to the very poor discrimination of the last option ("Respiratory system").

---

## Question 15.2

Explain why surgical instruments used in operating theatres are sterilised.

### Scoring and Comments on Question 15.2

#### Full Credit

Code 21: Student mentions both the need to ensure that there are no bacteria/germs on the instruments AND that this stops the spread of disease.

- To stop bacteria getting in the body and infecting the patient.
- So that no germs get into the body of another person going in for major surgery.

#### Partial Credit

Code 12: Student mentions the need to ensure that there are no bacteria, BUT not that this stops the spread of disease.

- To kill the germs on them.

Code 11: Student mentions that this stops the spread of disease, BUT not that it is because any bacteria on the instruments are killed.

- So the patient is not infected.
- To prevent any transfer of disease.

#### No Credit

Code 01: Other responses

- To keep them clean.
- Because patients are vulnerable during surgery.

Code 99: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Social



This item showed very good discrimination in the field trial and was of medium difficulty. Females were more likely to answer it correctly than were males.

Coders experienced difficulties in distinguishing between codes 11 and 12 in many cases and double-digit coding would not have been used if this item had been included in the main study.

---

### Question 15.3

Patients may be unable to eat and drink after surgery and so they are put on a drip (infusion) that contains water, sugars and mineral salts. Sometimes antibiotics and tranquillisers are also added to the drip.

Why are the sugars that are added to the drip important for the post-operative patient?

A. To avoid dehydration

B. To control post-operative pain

C. To cure post-operative infections

D. To provide necessary nutrition

### Scoring and Comments on Question 15.3

#### Full Credit

Code 1: D. To provide necessary nutrition

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Living systems (Knowledge of science)

**Application area:** Health

**Setting:** Personal / Social

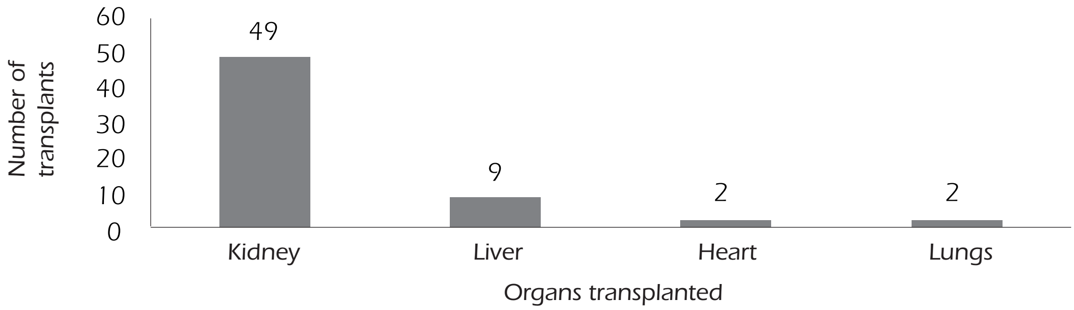
Field trial outcomes showed this item to be very easy and have an adequate discrimination. About 70% of students got it correct.

---



### Question 15.4

Organ transplants involve major surgery and are becoming more and more common. In the graph below, the numbers of transplants carried out in a particular hospital during 2003 are given.




Can the following conclusions be drawn *from the graph above?*

*Circle "Yes" or "No" for each conclusion.*

| Can this conclusion be drawn from the graph? | Yes or No? |
|---|---|
| If the lungs are transplanted, the heart must be transplanted too. | Yes / No |
| Kidneys are the most important organs in the human body. | Yes / No |
| Most of the patients that have a transplant have suffered from a kidney disease. | Yes / No |

### Scoring and Comments on Question 15.4

#### Full Credit

Code 1: All three correct: No, No, Yes in that order

#### No Credit

Code 0: Other responses

Code 9: Missing

**Item type:** Complex multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Health

**Setting:** Social

This item assesses students' skills in interpreting scientific data in graphical form and hence drawing appropriate conclusions. No specific information external to the item stimulus has to be drawn upon. Consequently, the appropriate knowledge classification is "Knowledge about science", category: "Scientific explanations".

The item performed very well in the field trial with good discrimination and medium difficulty.



### Question 15.5

How much interest do you have in the following information?

*Tick only one box in each row.*

| | High Interest | Medium Interest | Low Interest | No Interest |
|---|---|---|---|---|
| a) Learning how surgical instruments are sterilised | □ 1 | □ 2 | □ 3 | □ 4 |
| b) Knowing about the different types of anaesthetics that are used | □ 1 | □ 2 | □ 3 | □ 4 |
| c) Understanding how a patient's level of consciousness is monitored during surgery | □ 1 | □ 2 | □ 3 | □ 4 |

**Item type:** Attitudinal

**Attitude:** Interest in learning about science

This item was designed to assess a student's interest in aspects of the science of surgery. Like all attitudinal items, it comes at the end of the unit so that the students are familiar with the context prior to being asked their opinions.

Exploratory factor analysis of the field trial results showed all three statements loading somewhat on an "interest in science" dimension, but more heavily on another dimension thought to represent interest in/concern for health and safety. The level of interest expressed varied across the three statements, with most interest shown in the last statement and least interest shown in the first statement.

## SCIENCE UNIT 16

### Wind Farms

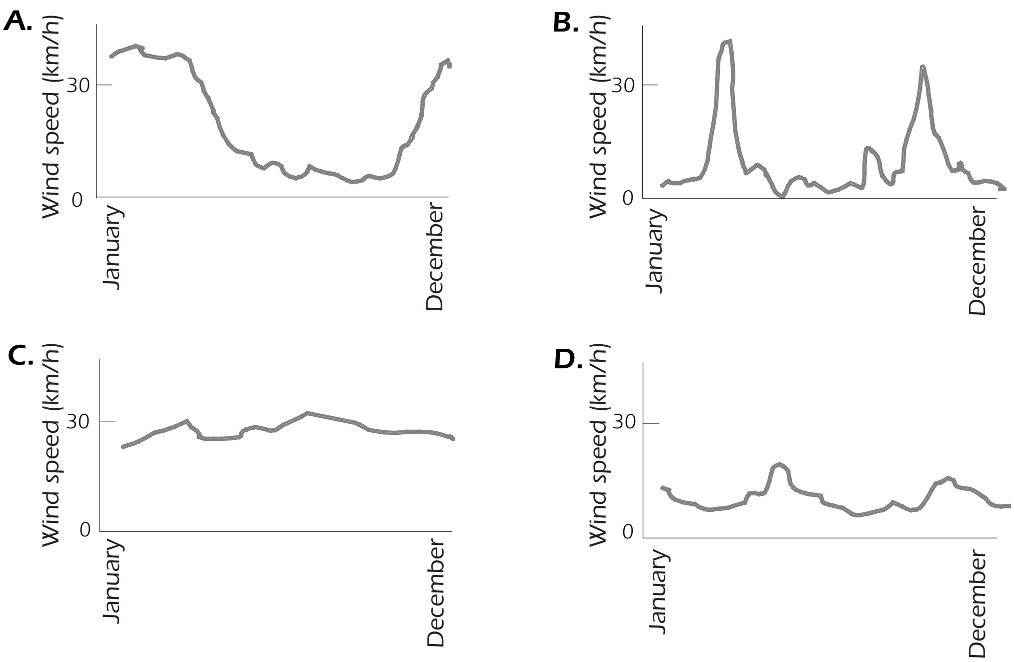
Many people believe that wind should replace oil and coal as a source of energy for producing electricity. The structures in the picture are windmills with blades that are rotated by the wind. These rotations cause electricity to be produced by generators that are turned by the windmills.

A wind farm

---

### Question 16.1

The graphs below show the average wind speeds in four different places throughout a year. Which one of the graphs indicates the most appropriate place to establish a wind farm for generating electricity?




### Scoring and Comments on Question 16.1

**Full Credit**

Code 1: C

**No Credit**

Code 0: Other responses

Code 9: Missing



**Item type:** Multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Technology systems (Knowledge of science) / Scientific explanations (Knowledge about science)

**Application area:** Natural resources

**Setting:** Social

Concerns about the use of fossil fuels to generate electric power feature regularly in the media. Potential and actual alternatives have impacts on the way people live and can generate their own environmental issues. This unit was given a very high priority for inclusion in the main study by country representatives.

Students need to bring to this item knowledge that the higher the wind speed the more electricity that is generated, and that there are advantages for distribution in consistency in wind speed. This suggests a classification of "Knowledge of science", category: "Technology systems". The graphical data must then be interpreted in light of that knowledge and this suggests a classification of "Knowledge about science", category: "Scientific explanations".

The version of this item used in the field trial proved to be very easy with 75% of students responding correctly. In the version shown here, the option C graph has been shifted down slightly and this may increase the item's difficulty a little.

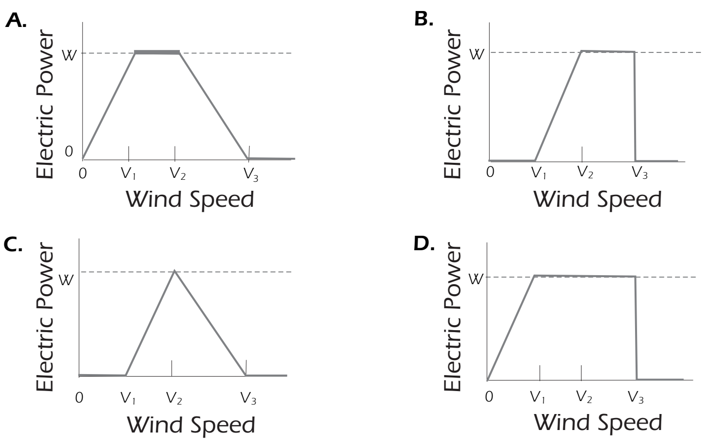
---

### Question 16.2

The stronger the wind, the faster the windmill blades rotate and the greater the electric power output. However, there is not a direct relationship between wind speed and electric power in a real setting. Below are four working conditions of electricity generation in a real wind farm.

- The windmill blades start rotating when the wind speed reaches $V_1$.
- The electric power output reaches a maximum (W) when the wind speed is $V_2$.
- For safety reasons, the blades are prevented from rotating faster than they do when the wind speed is $V_2$.
- The blades stop rotating when the wind speed reaches $V_3$.

Which one of the following graphs best represents the relationship between wind speed and electric power output under these working conditions?





### Scoring and Comments on Question 16.2

**Full Credit**

Code 1: B.

**No Credit**

Code 0: Other responses

Code 9: Missing

**Item type:** Multiple choice

**Competency:** Using scientific evidence

**Knowledge category:** Scientific explanations (Knowledge about science)

**Application area:** Natural resources

**Setting:** Social

Students must match a set of conditions to graphical features in interpreting the graphs provided. The conditions represent a set of technical data rather than experimental data. The item was not considered for inclusion in the main study because of concern that it was mainly testing mathematical literacy and because the opening sentence provided information helpful in answering the previous question.

The item was found to be of medium difficulty in the field trial with an adequate discrimination. However, in a number of countries the mean ability of the students who chose distractor C was not much below the mean ability of the students who chose the correct option (B).

---

## Question 16.3

The higher the altitude the more slowly windmills rotate at the same wind speed.

Which one of the following is the best reason why the blades of windmills rotate more slowly in higher places at the same wind speed?

A. The air is less dense as altitude increases.

B. The temperature is lower as altitude increases.

C. Gravity becomes less as altitude increases.

D. It rains more often as altitude increases.

### Scoring and Comments on Question 16.3

**Full Credit**

Code 1: A. The air is less dense as altitude increases.

**No Credit**

Code 0: Other responses

Code 9: Missing



**Item type:** Multiple choice

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Earth and space systems (Knowledge of science)

**Application area:** Natural resources

**Setting:** Social

Like all items in the unit, this item was regarded highly for inclusion in the main study by country representatives. However, the field trial results indicated some issues which led to its rejection. In particular, there was some variability in item difficulty between countries and distractor D was found to be rather weak. Males were more likely to answer the item correctly than were females.

---

## Question 16.4

Describe one specific advantage, and one specific disadvantage, of using wind to generate electricity compared with using fossil fuels like coal and oil.

An advantage ..................................................

..................................................

A disadvantage ..................................................

..................................................

### Scoring and Comments on Question 16.4

#### Full Credit

Code 2: One **specific** advantage and one **specific** disadvantage are described.

**Scoring Comment:** It is possible for the cost of wind farms to be seen as an advantage or disadvantage depending on what aspect is considered (e.g. establishment costs or running costs). Hence, mentioning the cost involved, without further explanation, is not sufficient to gain credit as either an advantage or a disadvantage.

[Advantage]

- Do not discharge carbon dioxide ($CO_2$).
- Do not consume fossil fuels.
- The wind resource will not be used up.
- After the wind generator is established, the cost for electric generation is cheap.
- No waste and/or no toxic substance will be emitted.
- Using natural forces or clean energy.
- Environmentally friendly and will last for a very long time.

[Disadvantage]

- Generation on demand is not possible. [Because the wind speed cannot be controlled.]
- Good places for windmills are limited.
- The windmill could be damaged by too strong wind.
- The amount of power generated by each windmill is relatively small.
- Noise pollution occurs in some cases.
- Birds are sometimes killed when they crash into the rotors.
- Natural views are altered. [Visual pollution.]
- Expensive to set up.



### Partial Credit

Code 1: Either a correct advantage or a correct disadvantage is described (as shown in the full credit examples) but not both.

### No Credit

Code 0: No correct advantage or correct disadvantage is described. Individual examples of unacceptable advantages or disadvantages are given below.

- *Good for the environment or nature. [This answer is a general value statement.]*
- *Bad for the environment or nature.*
- *It costs less to build a wind power generator than to build a fossil fuel power plant. [This ignores the fact that a great number of wind power generators would be needed to produce the same amount of power as a fossil fuel power plant.]*
- *It wouldn't cost as much.*

Code 9: Missing

**Item type:** Open-constructed response

**Competency:** Explaining phenomena scientifically

**Knowledge category:** Technology systems (Knowledge of science)

**Application area:** Natural resources

**Setting:** Social

This item is open to a wide range of correct or partially correct responses and this led to coding difficulties in the field trial. Most of these difficulties concerned references to cost and so a scoring comment has been added to this version to clarify how such responses should be assessed.

---

# Annex B

## PISA EXPERT GROUPS



## Science Expert Group

*Chair*
*Rodger Bybee*
*Biological Sciences Curriculum Study*
*Colorado Springs, United States*

Ewa Bartnik
University of Warsaw
Warsaw, Poland

Peter Fensham
Monash University
Queensland, Australia

Paulina Korsnakova
Department of Educational Measurement
Bratislava, Slovak Republic

Robert Laurie
Department of Education of New Brunswick
New Brunswick, Canada

Svein Lie
University of Oslo
Blindern, Norway

Pierre Malléus
Ministry for National Education, Higher Education and Research
Champigneulles, France

Michelina Mayer
National Insitute for the Evaluation of Instructional Systems
Rome, Italy

Robin Millar
University of York
York, United Kingdom

Yasushi Ogura
National Institute for Educational Policy Research
Tokyo, Japan

Manfred Prenzel
University of Kiel
Kiel, Germany

Andrée Tiberghien
University of Lyon
Ste Foy les Lyon, France

## Reading Expert Group

*Chair*
*John de Jong*
*Language Testing Services*
*Oranjestraat, Netherlands*

Irwin Kirsch
Educational Testing Service
Princeton, New Jersey, United States

Marilyn Binkley
National Center for Educational Statistics
Washington, D.C., United States

Alan Davies
University of Edinburgh
Scotland, United Kingdom

Stan Jones
Statistics Canada
Nova Scotia, Canada

Dominique Lafontaine
University of Liège
Liège, Belgium

Pirjo Linnakylä
University of Jyväskylä
Jyväskylä, Finland

Martine Rémond
IUFM de Créteil
University of Paris 8
Andresy, France



**Mathematics Expert Group**

*Chair*
*Jan de Lange*
*Utrecht University*
*Utrecht, Netherlands*

Werner Blum
University of Kassel
Kassel, Germany

John Dossey
Illinois State University
Eureka, Illinois, United States

Zbigniew Marciniak
Warsaw University
Warsaw, Poland

Mogens Niss
IMFUFA, Roskilde University
Roskilde, Denmark

Yoshinori Shimizu
University of Tsukuba
Tsukuba-shi, Ibaraki, Japan

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(98 2006 03 1 P) ISBN 92-64-02639-8 No. 55285 2006

# Assessing Scientific, Reading and Mathematical Literacy
## A FRAMEWORK FOR PISA 2006

Are students well prepared to meet the challenges of the future? Are they able to analyse, reason and communicate their ideas effectively? Do they have the capacity to continue learning throughout life?

These are the questions that the OECD's Programme for International Student Assessment (PISA) aims to answer through three-yearly surveys (2000, 2003, 2006...) of 15-year-olds in the principal industrialised countries that examine student performance in reading, mathematical and scientific literacy. The first survey, in 2000, focused on reading and the second, in 2003, on mathematics. PISA 2006, with its focus on science, completes the first full cycle of the survey.

*Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* presents the conceptual framework underlying the PISA 2006 survey. It includes a re-developed and expanded framework for scientific literacy, an innovative component on the assessment of students' attitudes to science and the frameworks for the assessment of reading and mathematics. Within each domain, the framework defines the content that students need to acquire, the processes that need to be performed, and the contexts in which knowledge and skills are applied. The domains and their aspects are also illustrated with sample tasks.

**FURTHER READING**

The first results from PISA 2003 were published in *Learning for Tomorrow's World – First Results from PISA 2003* (OECD, 2004) and *Problem Solving for Tomorrow's World – First Measures of Cross-Curricular Competencies from PISA 2003* (OECD, 2004).

***The OECD Programme for International Student Assessment (PISA)***

*PISA is a collaborative process among the 30 member countries of the OECD and nearly 30 partner countries. It brings together scientific expertise from the participating countries and is steered by their governments on the basis of shared, policy-driven interests. PISA is an unprecedented attempt to measure student achievement, as is evident from some of its features:*

– *The literacy approach: PISA aims to define each assessment area (science, reading and mathematics) not mainly in terms of mastery of the school curriculum, but in terms of the knowledge and skills needed for full participation in society.*

– *A long-term commitment: It will enable countries to monitor regularly and predictably their progress in meeting key learning objectives.*

– *The age-group covered: By assessing 15-year-olds,* i.e. *young people near the end of their compulsory education, PISA provides a significant indication of the overall performance of school systems.*

– *The relevance to lifelong learning: PISA does not limit itself to assessing students' knowledge and skills but also asks them to report on their own motivation to learn, their beliefs about themselves and their learning strategies.*

The full text of this book is available on line via this link:
*http://www.sourceoecd.org/education/9264026398*

Those with access to all OECD books on line should use this link:
*http://www.sourceoecd.org/9264026398*

SourceOECD is the OECD's online library of books, periodicals and statistical databases. For more information about this award-winning service and free trials ask your librarian, or write to us at **SourceOECD@oecd.org**.

www.pisa.oecd.org

OECD PUBLISHING

ISBN 92-64-02639-8
98 2006 03 1 P