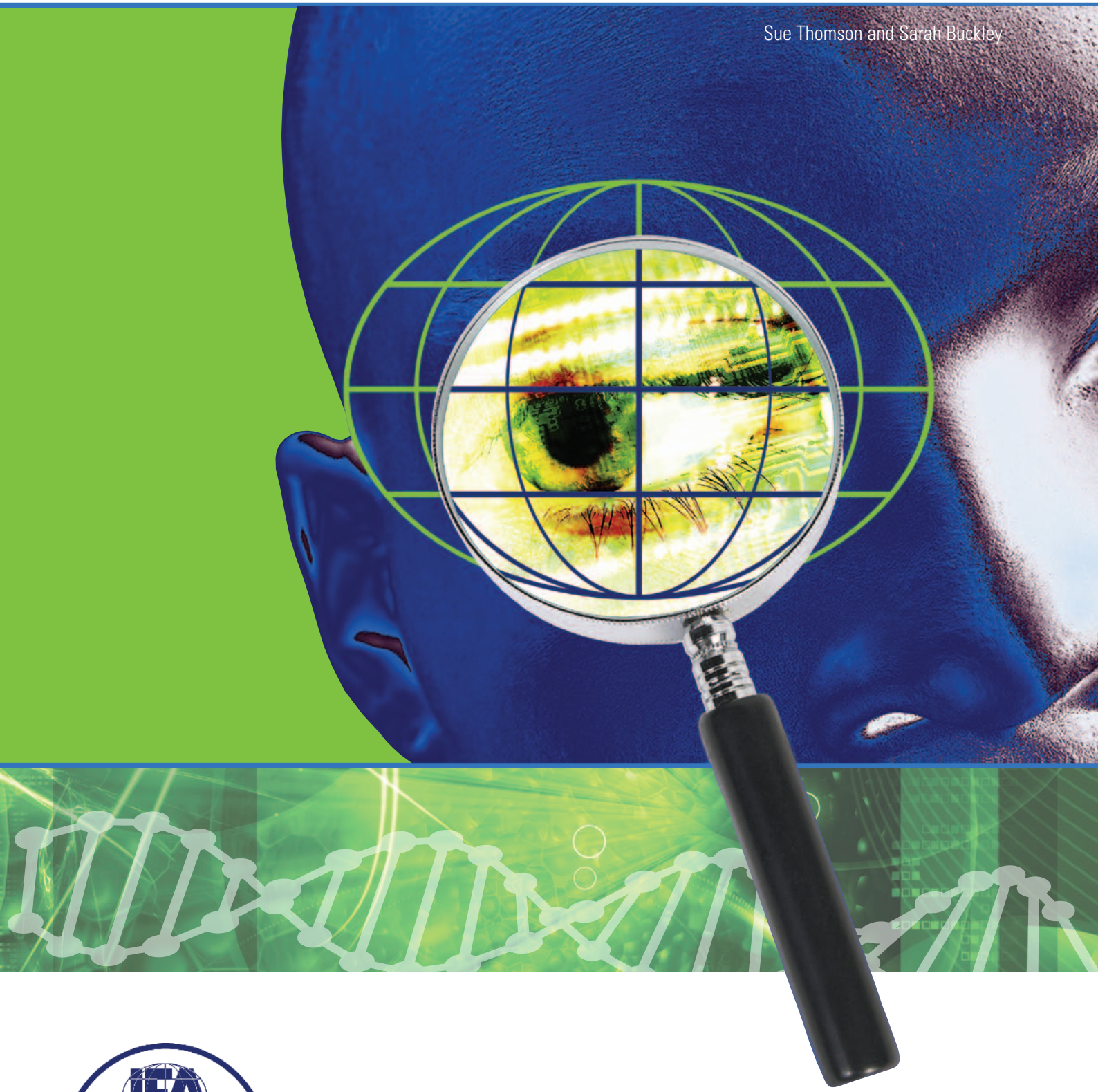
Informing science pedagogy: TIMSS 2007

Australia and the world

Sue Thomson and Sarah Buckley



A further investigation from the Trends in International Mathematics and Science Study (TIMSS) 2007

National Library of Australia Cataloguing-in-Publication entry

Author:	Thomson, Sue.
Title:	Informing science pedagogy : TIMSS 2007 Australia and the world / Sue Thomson, Sarah Buckley.
ISBN:	9780864318299 (pbk.)
Subjects:	Trends in International Mathematics and Science Study. Science--Study and teaching--Australia
Other Authors/ Contributors:	Buckley, Sarah. Australian Council for Educational Research.
Dewey Number:	507.10994

Contents

This report

This report aims to provide teachers with more detailed information on Australian students' capabilities in science in terms of the TIMSS assessment. A detailed assessment of Australian students' performance can be found in the full report that was written to inform the educational community about Australian students' performance on national and international scales. The report - *TIMSS 07: Taking a closer look at mathematics and science in Australia* - and more information about TIMSS can be accessed at www.acer.edu.au/timss.

What is TIMSS?

TIMSS is a key part of MCEETYA's National Assessment Program.

TIMSS is the Trends in International Mathematics and Science Study. TIMSS 2007 was the fourth in a cycle of internationally comparative assessments, conducted by the International Association for the Evaluation of Educational Achievement (IEA). The IEA is dedicated to improving teaching and learning in mathematics and science for students around the world.

Carried out every four years with Year 4 and Year 8 students, TIMSS provides data about national and international trends in mathematics and science achievement. In Australia, TIMSS is part of the Ministerial Council on Education, Employment, Training and Youth Affairs' (MCEETYA) National Assessment Program. It provides a level of international benchmarking to complement national assessments at Years 3,5,7 and 9 and other sample-based national studies.

To inform educational policy in the participating countries, TIMSS also routinely collects extensive background information that addresses concerns about the quantity, quality, and content of instruction.

What is the focus of TIMSS?

The main goal of TIMSS is to assist countries to monitor and evaluate their mathematics and science teaching across time and across year levels.

TIMSS has a curriculum focus. The three levels of the curriculum defined by TIMSS are:

TIMSS examines three levels of the curriculum: what is intended to be taught, what is actually taught, and what it is that students learn!

The *intended* curriculum – the curriculum as specified at national or system level.

- What are mathematics and science students around the world expected to learn?
- How do countries vary in their intended goals, and
- What characteristics of education systems, schools and students influence the development of these goals?

The *implemented* curriculum – the curriculum as interpreted and delivered by classroom teachers.

- What opportunities are provided for students to learn mathematics and science?
- How do instructional practices vary among countries, and
- What factors influence these variations?

The *attained* curriculum – which is that part of the curriculum that is learned by students, as demonstrated by their attitudes and achievements.

- What mathematics and science concepts, processes and attitudes have students learned?
- What factors are linked to students' opportunity to learn, and
- How do these factors influence students' achievements?

What do participants have to do?

Students complete an assessment booklet that contains an equal number of questions about mathematics and science. Questions are presented in two general formats: multiple choice and constructed response. After the students complete the assessment booklet they complete a short questionnaire. Teachers and principals also complete questionnaires.

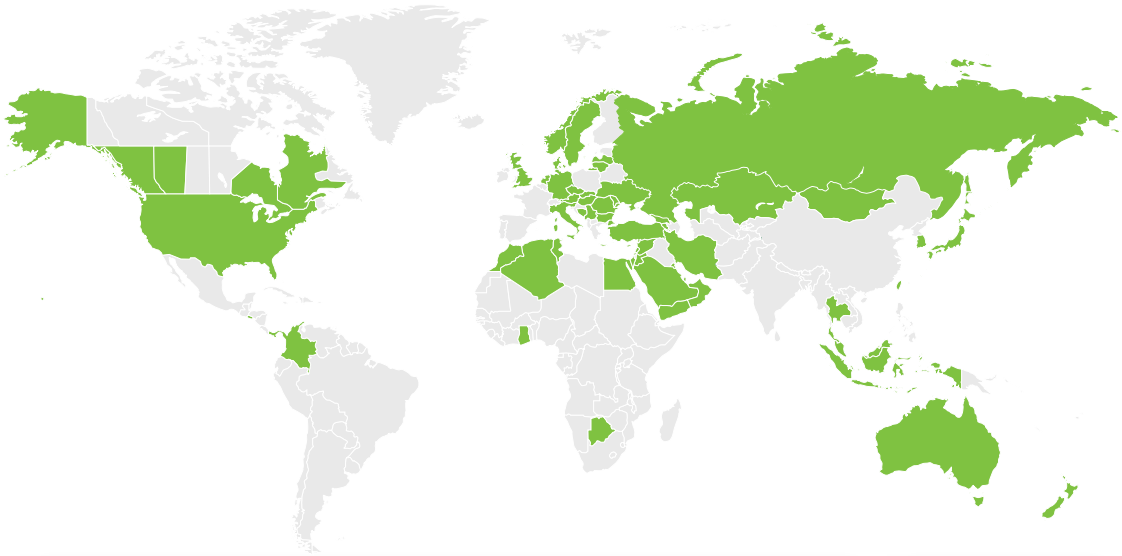
These internationally standard questionnaires gather information at the student, class, and school level. The student questionnaire gathers information from students about their family background, aspects of learning and instruction in mathematics and science, and the context of instruction. The teacher questionnaire collects information about a variety of issues related to qualifications, pedagogical practices, teaching styles, use of technology, assessment and assignment of homework, and classroom climate. The school questionnaire, answered by the principal, gathers descriptive information about the school and information about instructional practices. For example, questions were asked about recruitment of teachers and numbers of staff, teacher morale, school and teacher autonomy, school resources, and school policies and practices such as the use of student assessments.

Students complete an assessment and a background questionnaire. Teachers and principals also complete a questionnaire.

Who participated in TIMSS 2007?

Internationally

A total of 49 countries at Year 8 and 36 countries at Year 4 participated in TIMSS 2007. In addition, four provinces of Canada, two states of the United States, Dubai Emirate, UAE and Basque Country, Spain, were also in the study as benchmarking participants*. These are shown in Figure 1.



Participating Countries

Algeria	Cyprus	Hungary	Latvia	Oman	Slovenia
Armenia	Czech Republic	Indonesia	Lebanon	Palestinian Nat'l Auth.	Sweden
Australia	Denmark	Iran, Islamic Rep. of	Lithuania	Qatar	Syrian Arab Republic
Austria	Egypt	Israel	Malaysia	Romania	Thailand
Bahrain	El Salvador	Italy	Malta	Russian Federation	Tunisia
Bosnia & Herzegovina	England	Japan	Mongolia	Saudi Arabia	Turkey
Botswana	Georgia	Jordan	Morocco	Scotland	Ukraine
Bulgaria	Germany	Kazakhstan	Netherlands	Serbia	United States
Chinese Taipei	Ghana	Korea, Rep. of	New Zealand	Singapore	Yemen
Colombia	Hong Kong SAR	Kuwait	Norway	Slovak Republic	

Benchmarking Participants

Alberta, Canada
Basque Country, Spain
British Colombia, Canada
Dubai, UAE
Massachusetts, US
Minnesota, US
Ontario, Canada
Quebec, Canada

Figure 1 Map of participating countries

* Benchmarking participants are provinces or regions that participated in TIMSS for their own internal benchmarking purposes. Data from these regions are not included in the international average.

In Australia

Schools are chosen from all schools in Australia to be representative of their state and sector.

A stratified random sample of 230 primary schools and 230 secondary schools was chosen in Australia, and of this sample, 229 primary schools and 228 secondary schools participated in the data collection for TIMSS 2007. The sample is drawn from all schools in Australia, and is representative of all states and sectors. In each state, government, Catholic and independent schools are chosen proportional to their number in the state. Figure 2 shows all schools in Australia (including those on Christmas Island, Norfolk Island and King Island) in green and all schools selected for TIMSS in black.

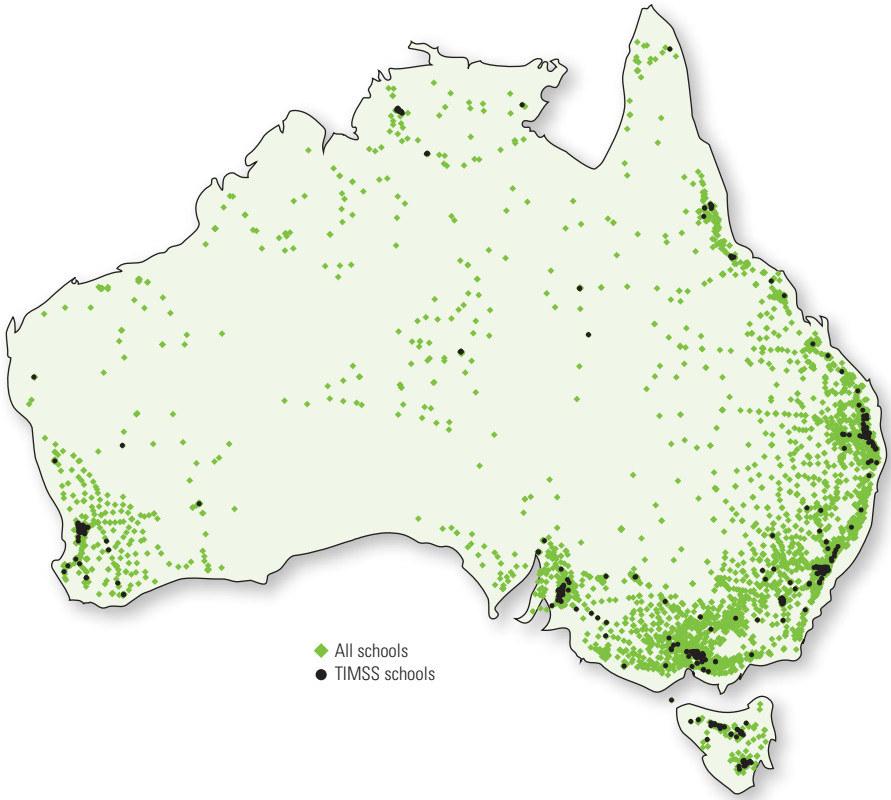


Figure 2 Australian schools and TIMSS sample schools

How is science assessed in TIMSS?

TIMSS is curriculum based. All countries are surveyed to ensure that the test is kept up-to-date.

At the beginning of each TIMSS cycle, a committee comprised of curriculum experts in mathematics and science from a range of participating countries reviews the framework for the upcoming assessment. This framework is finally ratified by all member countries as being representative of their country's curricula. For TIMSS to not become a lowest common denominator assessment, however, there will always be some content that is not covered in the curriculum of each country. This is managed at the data analysis stage of the project by removing the items that a country argues is outside their curriculum. It rarely makes any difference to a country's score.

The assessment has both a content dimension and a cognitive dimension.

The general framework for the TIMSS 2007 assessment of science had two dimensions, one relating to context and the other relating to cognition. Within the content dimension, there were three domains in science at Year 4 and four at Year 8. In addition there were three cognitive domains in each curriculum area: *knowing*, *applying* and *reasoning*. The two dimensions and their domains were the foundation of the science assessments. The content domains defined the

specific subject matter covered by the assessment, and the cognitive domains defined the sets of skills expected of students as they engage with the content. An elaboration of the content domains is shown below in Table 1 and Table 2, and includes proportions of each topic area examined in the TIMSS tests.

Table 1 TIMSS content domains in science at Year 4

	Science Year 4
Content Domains	**Topic areas**
Life science (45%)	Characteristics and life processes of living things
	Life cycles, reproduction and heredity
	Interaction with the environment
	Ecosystems
	Human health
Physical science (35%)	Classification and properties of matter
	Physical states and changes in matter
	Energy sources, heat, and temperature
	Light and sound
	Electricity and magnetism
	Forces and motion
Earth science (20%)	Earth's structure, physical characteristics, and resources
	Earth's processes, cycles and history
	Earth in the solar system

Table 2 TIMSS content domains in science at Year 8

	Science Year 8
Content Domains	**Topic areas**
Biology (35%)	Characteristics, classification, and life processes of organisms
	Cells and their functions
	Life cycles, reproduction, and heredity
	Diversity, adaptation, and natural selection
	Ecosystems
	Human health
Chemistry (20%)	Classification and composition of matter
	Properties of matter
	Chemical change
Physics (25%)	Physical states and changes in matter
	Energy transformations, heat, and temperature
	Light
	Sound
	Electricity and magnetism
	Forces and motion
Earth Science (20%)	Earth's structure and physical features
	Earth's processes, cycles and history
	Earth's resources, their use and conservation
	Earth in the solar system and the universe

The structure of the TIMSS assessment

A consequence of the assessment goals of TIMSS is that there are many more questions on the assessment than can be answered by a student in the amount of testing time available. To work around this, TIMSS uses an approach that involves packaging the entire assessment pool of mathematics and science questions into a set of 14 student achievement booklets, with each student completing just one booklet.

There are 14 different student achievement test booklets – these all form part of the jigsaw that we put together to enable the sample to represent all Australian students.

Each question, or item, appears in two booklets, providing a mechanism for linking together the student responses from the various booklets. Booklets are distributed among students in participating classrooms so that the groups of students completing each booklet are approximately equivalent in terms of student ability.

Using Item-Response Theory (IRT) scaling techniques, a comprehensive picture of the achievement of the entire student population is assembled from the combined responses of individual students to the booklets they are assigned. This approach reduces to manageable proportions what otherwise would be an impossible student burden, albeit at the cost of greater complexity in booklet assembly, data collection, and data analysis.

A number of new items are developed for each new TIMSS assessment. In addition, because TIMSS is an assessment that examines trends, a number of items are retained from one cycle to the next, to link assessments to each other.

Question types and scoring the responses

Two question formats are used in the TIMSS assessment – multiple-choice and constructed-response. At least half of the total number of points represented by all the questions will come from multiple-choice questions. Each multiple-choice question is worth one score point.

Multiple-Choice Questions. Multiple-choice questions provide four response options, of which only one is correct. These questions can be used to assess any of the behaviours in the cognitive domains. However, because they do not allow for students' explanations or supporting statements, multiple-choice questions may be less suitable for assessing students' ability to make more complex interpretations or evaluations.

In assessing Year 4 and Year 8 students, it is important that linguistic features of the questions be developmentally appropriate. Therefore, the questions are written clearly and concisely. The response options also are written succinctly in order to minimise the reading load of the question. The options that are incorrect are written to be plausible, but not deceptive. For students who may be unfamiliar with this test question format, the instructions given at the beginning of the test include a sample multiple-choice item that illustrates how to select and mark an answer.

Constructed-Response Questions. For this type of test item students are required to construct a written response, rather than select a response from a set of options. Constructed-response questions are particularly well-suited for assessing aspects of knowledge and skills that require students to explain phenomena or interpret data based on their background knowledge and experience.

The scoring guide for each constructed-response question describes the essential features of appropriate and complete responses. The guides focus on evidence of the type of behaviour the question assesses. They describe evidence of partially correct and completely correct responses. In addition, sample student responses at each level of understanding provide important guidance to those who will be rating the students' responses. In scoring students' responses to constructed-response questions, the focus is solely on students' achievement with respect to the topic being assessed, not on their ability to write well. However, students need to communicate in a manner that will be clear to those scoring their responses.

How results are reported in TIMSS

TIMSS summarises achievement for each year level in two ways. Firstly, results are reported on a scale with a mean of 500 and a standard deviation of 100. However, it should be noted that the results for year 4 and year 8 are not directly comparable. While the scales for the two year levels are expressed in the same numerical units, they are not directly comparable in terms of being able to say how much achievement or learning at one year level equals how much achievement or learning is observed in the other year level. That is, achievement cannot be described at either year level in absolute terms.

International Benchmarks

Another way in which general achievement was assessed in TIMSS was by examining the percentage of students in each country that reached certain benchmarks. While the achievement scales mentioned in the previous section summarise student performance in the cognitive and content knowledge measured by the TIMSS science tests, the international benchmarks help put these scores in context.

Internationally, it was decided that performance should be measured at four levels. These four levels summarise the achievement reached at:

- the 'advanced international benchmark', which was set at a score of 625;
- the 'high international benchmark', which was set at a score of 550;
- the 'intermediate international benchmark', which was set at a score of 475; and
- the 'low international benchmark', which was set at a score of 400.

The benchmarks discussed in this report were based solely on student performance in TIMSS 2007. It should also be noted that when reporting the percentage of students achieving a particular benchmark, this includes students achieving the benchmarks above this. For example, if 24 per cent of Year 8 students achieved the high international benchmark this would include the six per cent at the advanced benchmark.

Year 4 Science – Descriptors of performance at the international benchmarks

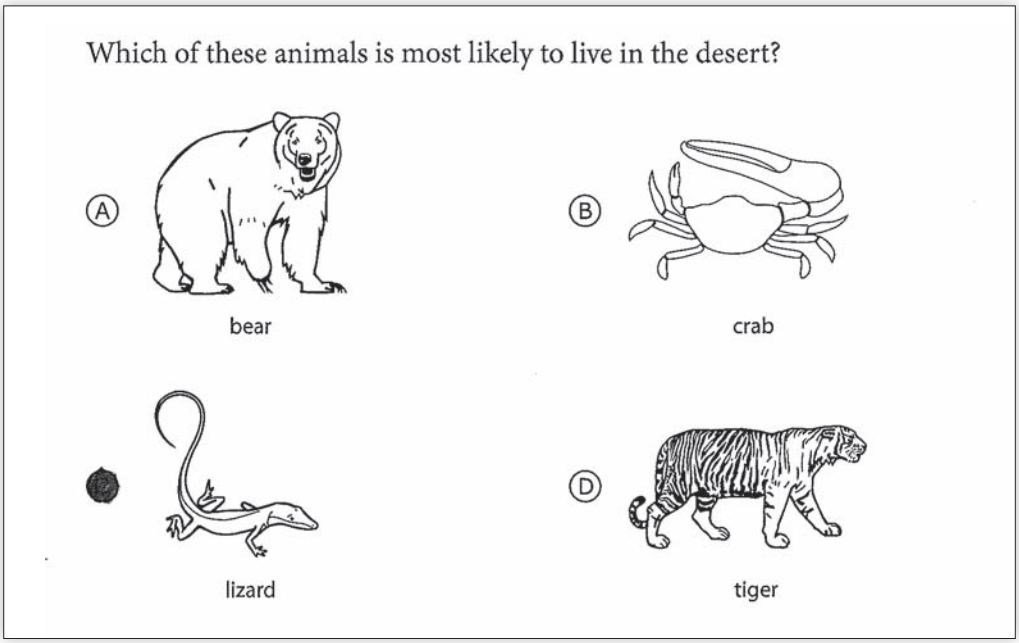
In Year 4 science, students at the Advanced International Benchmark were able to apply knowledge and understanding of scientific processes and relationships in beginning scientific inquiry, whereas those at the Low International Benchmark displayed some elementary knowledge of life science and physical science. Table 3 gives some brief descriptors of achievement at the international benchmarks for Year 4 science, and following this, an example is provided for each of the levels.

Table 3 Descriptors for Year 4 science international benchmarks

Year 4	Low International Benchmark	Intermediate International Benchmark	High International Benchmark	Advanced International Benchmark
	(400)	(475)	(550)	(625)
	Students have some elementary knowledge of life science and physical science. Students can demonstrate knowledge of some simple facts related to human health and the behavioural and physical characteristics of animals. They recognise some properties of matter, and demonstrate a beginning understanding of forces. Students interpret labelled pictures and simple diagrams, complete simple tables, and provide short written responses to questions requiring factual information.	*Students can apply basic knowledge and understanding to practical situations in the sciences.* Students recognise some basic information related to characteristics of living things and their interaction with the environment, and show some understanding of human biology and health. They also show some understanding of familiar physical phenomena. Students know some basic facts about the solar system and have a developing understanding of Earth's resources. They demonstrate some ability to interpret information in pictorial diagrams and apply factual knowledge to practical situations.	*Students can apply knowledge and understanding to explain everyday phenomena.* Students demonstrate some understanding of plant and animal structure, life processes, and the environment and some knowledge of properties of matter and physical phenomena. They show some knowledge of the solar system, and of Earth's structure, processes, and resources. Students demonstrate beginning scientific inquiry knowledge and skills, and provide brief descriptive responses combining knowledge of science concepts with information from everyday experience of physical and life processes.	*Students can apply knowledge and understanding of scientific processes and relationships in beginning scientific inquiry.* Students communicate their understanding of characteristics and life processes of organisms as well as of factors relating to human health. They demonstrate understanding of relationships among various physical properties of common materials and have some practical knowledge of electricity. Students demonstrate some understanding of the solar system and Earth's physical features and processes. They show a developing ability to interpret the results of investigations and draw conclusions as well as a beginning ability to evaluate and support an argument.

Year 4 Science: Performance at the Low International Benchmark

The low international benchmark for science included knowing some simple facts about human health and the behavioural and physical characteristics of animals and humans. The item shown below presents an example of student achievement at this benchmark. Students were presented with a pictorial representation of four animals and asked to identify the animal most likely to live in the desert. On average internationally 68 per cent of Year 4 students were able to identify the lizard as the most likely desert dweller. More than 90 per cent of students in the United States correctly answered this item, and 88 per cent of Australian students also identified the correct animal. This was significantly higher than the international average.



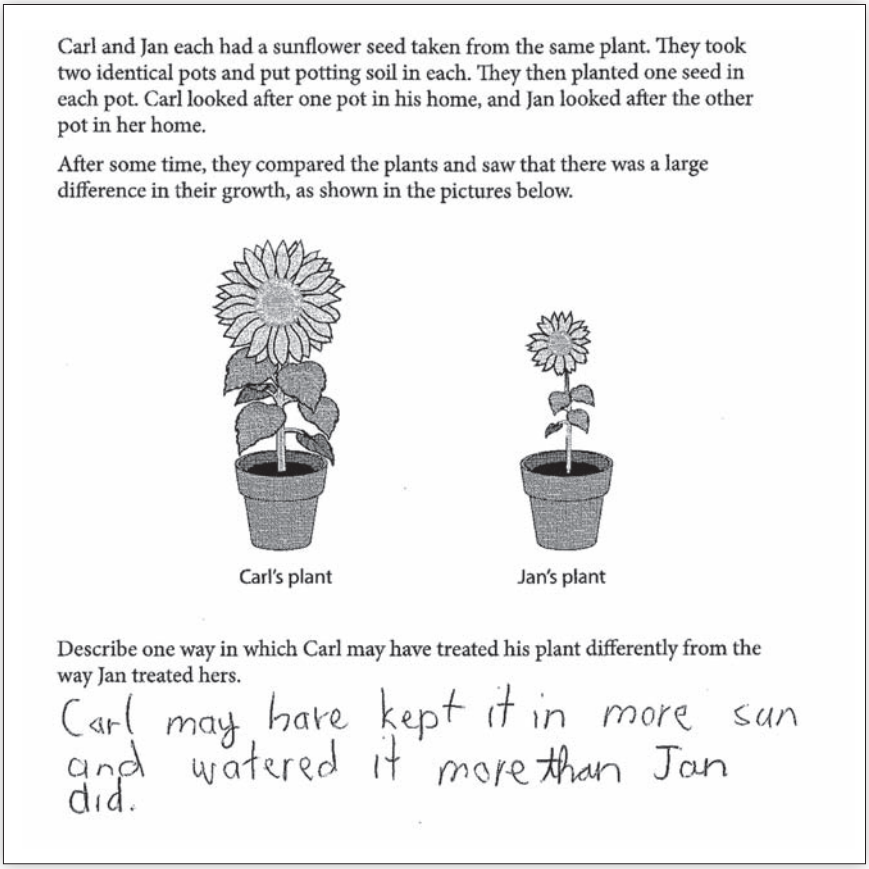
Country	Per cent full correct
United States	92 (0.8)
Russian Federation	89 (2.2)
Australia	88 (1.8)
Kazakhstan	86 (2.3)
England	84 (1.9)
Denmark	84 (2.1)
New Zealand	81 (1.4)
International average	**68 (0.4)**
Yemen	28 (2.1)

Year 4 Science: Performance at the Intermediate International Benchmark

Students achieving at the intermediate international benchmark were able to apply basic knowledge and understanding to practical situations in the sciences. For example they could recognise some basic information about characteristics of living things.

In the following example, students were shown a picture of two sunflower plants grown in similar pots of soil from seeds from the same plant. One plant was clearly larger and healthier looking than the other, and to earn full credit on this item students had to describe one way in which the larger plant may have been treated differently from the smaller one.

On average across countries 63 per cent of students answered correctly, explaining, for example, that the larger plant may have been given more light and water. Students in Australia and England did very well on this item, with 83 and 81 per cent of students, respectively, gaining credit.



Carl and Jan each had a sunflower seed taken from the same plant. They took two identical pots and put potting soil in each. They then planted one seed in each pot. Carl looked after one pot in his home, and Jan looked after the other pot in her home.

After some time, they compared the plants and saw that there was a large difference in their growth, as shown in the pictures below.

Carl's plant

Jan's plant

Describe one way in which Carl may have treated his plant differently from the way Jan treated hers.

Carl may have kept it in more sun and watered it more than Jan did.

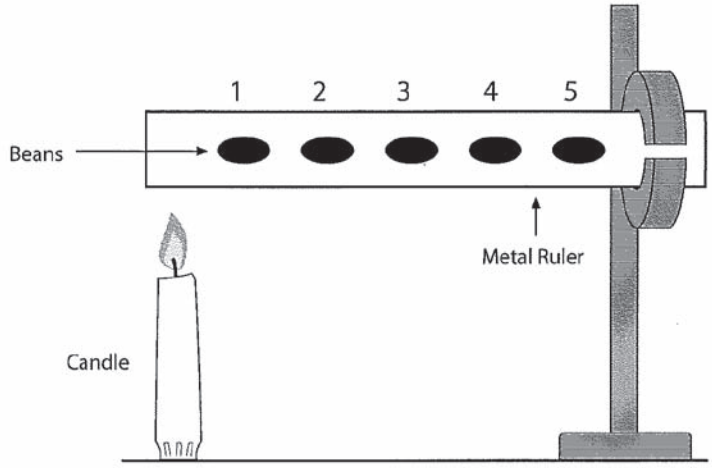
Country	Per cent full correct
Singapore	85 (1.8)
Lithuania	85 (1.7)
Netherlands	84 (2.0)
Australia	83 (2.3)
England	81 (1.8)
United States	78 (1.2)
New Zealand	77 (1.6)
International average	**63 (0.4)**
Yemen	7 (1.2)

Year 4 Science: Performance at the High International Benchmark

Students achieving the high international benchmark in science at Year 4 demonstrated some competency with many of the topics in the framework. At this level, students demonstrate some knowledge of life processes, and some knowledge of properties of matter and physical phenomena.

In this example, students were required to recognise that when heat is applied to one end of a metal ruler, that heat will be conducted to the other end.

On average internationally this was answered correctly by 57 per cent of Year 4 students. In Japan 92 per cent of students were able to answer correctly, and in both England and the United States around two-thirds of students answered correctly. In Australia 59 per cent of students answered correctly.



Beans are fixed on a metal ruler with butter as shown in the figure above. The ruler is heated at one end. In which order will the beans fall off?

- ● 1, 2, 3, 4, 5
- Ⓑ 5, 4, 3, 2, 1
- Ⓒ 1, 3, 5, 4, 2
- Ⓓ All at the same time

Country	Per cent full correct
Japan	92 (1.2)
Singapore	88 (1.4)
Hong Kong	75 (2.1)
England	67 (2.3)
United States	66 (1.7)
Australia	59 (2.8)
New Zealand	58 (2.2)
International average	**57 (0.4)**
Yemen	20 (1.6)

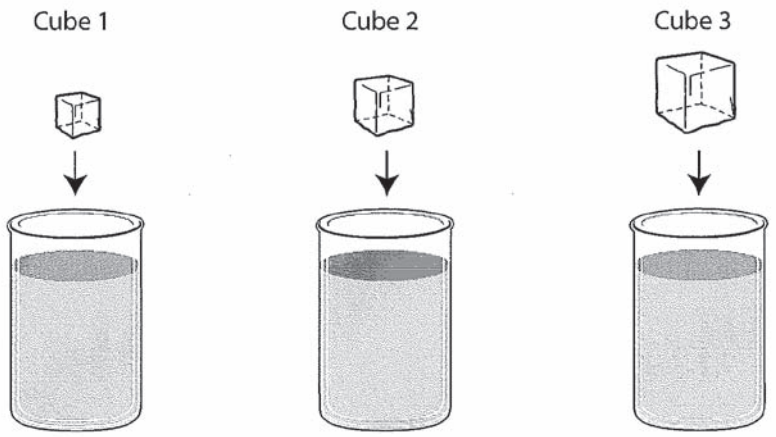
Year 4 Science: Performance at the Advanced International Benchmark

Students achieving at or above the advanced benchmark at Year 4 demonstrated fluency with most framework topics. They also successfully completed items at the high, intermediate and low benchmarks.

This item provides an example of a question that students achieving at this benchmark would be expected to have answered correctly. While there was no expectation that the concept of density will be mastered, students at the fourth grade were expected to understand that an object's capacity to sink or float is not determined by its size. Students were presented with a diagram showing three beakers the same size and containing the same amount of water, and three ice cubes of varying size.

On average across countries 39 per cent of students recognised that all three ice cubes would float, regardless of their size. Sixty per cent of the students in Chinese Taipei answered this item correctly, closely followed by Japan, Singapore, Austria and Australia.

Susie has three ice cubes of different sizes. She places each ice cube into an identical beaker containing the same volume of water, as shown in the diagram.



What happens to the ice cubes when they are placed in the water?

Ⓐ Cubes 1, 2, and 3 will sink.

Ⓑ Cubes 1, 2, and 3 will float.

Ⓒ Cube 1 will float, and cubes 2 and 3 will sink.

Ⓓ Cubes 1 and 2 will float, and cube 3 will sink.

Country	Per cent full correct
Chinese Taipei	60 (2.1)
Japan	58 (2.3)
Singapore	57 (1.8)
Australia	56 (2.4)
United States	44 (1.6)
New Zealand	44 (2.5)
International average	**39 (0.4)**
England	37 (2.2)
Algeria	20 (1.8)

Year 8 Science – Descriptors of performance at the international benchmarks

In Year 8 science, students at the Advanced International Benchmark demonstrated a grasp of some complex and abstract concepts in biology, chemistry, physics, and Earth science. In comparison, those at the Low International Benchmark simply recognised some basic facts from the life and physical sciences.

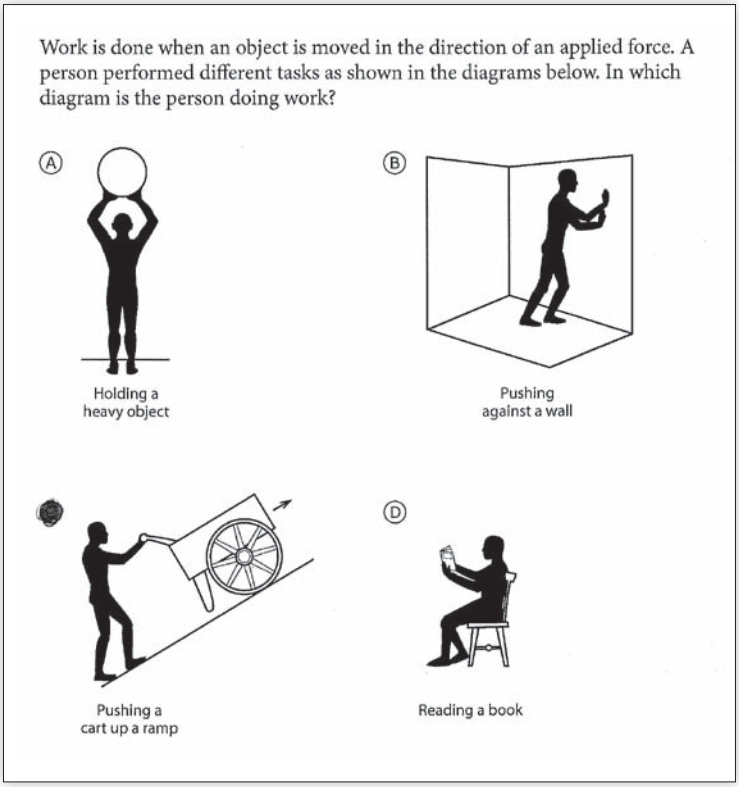
Table 4 provides descriptors for each of the benchmarks at Year 8 level, and this table is followed by examples of each level of the benchmark.

Table 4 Descriptors for Year 8 science international benchmarks

Year 8	Low International Benchmark	Intermediate International Benchmark	High International Benchmark	Advanced International Benchmark
	(400)	(475)	(550)	(625)
	Students can recognize some basic facts from the life and physical sciences. They have some knowledge of the human body, and demonstrate some familiarity with everyday physical phenomena. Students can interpret pictorial diagrams and apply knowledge of simple physical concepts to practical situations.	*Students can recognise and communicate basic scientific knowledge across a range of topics.* They demonstrate some understanding of characteristics of animals, food webs, and the effect of population changes in ecosystems. They are acquainted with some aspects of sound and force and have elementary knowledge of chemical change. They demonstrate elementary knowledge of the solar system, Earth's processes, and resources and the environment. Students extract information from tables and interpret pictorial diagrams. They can apply knowledge to practical situations and communicate their knowledge through brief descriptive responses.	*Students can demonstrate conceptual understanding of some science cycles, systems, and principles.* They have some understanding of biological concepts including cell processes, human biology and health, and the interrelationship of plants and animals in ecosystems. They apply knowledge to situations related to light and sound, demonstrate elementary knowledge of heat and forces, and show some evidence of understanding the structure of matter, and chemical and physical properties and changes. They demonstrate some understanding of the solar system, Earth's processes and resources, and some basic understanding of major environmental issues. Students demonstrate some scientific inquiry skills. They combine information to draw conclusions, interpret tabular and graphical information, and provide short explanations conveying scientific knowledge.	*Students can demonstrate a grasp of some complex and abstract concepts in biology, chemistry, physics, and Earth science.* They have an understanding of the complexity of living organisms and how they relate to their environment. They show understanding of the properties of magnets, sound, and light, as well as demonstrating understanding of structure of matter and physical and chemical properties and changes. Students apply knowledge of the solar system and of Earth's features and processes, and apply understanding of major environmental issues. They understand some fundamentals of scientific investigation and can apply basic physical principles to solve some quantitative problems. They can provide written explanations to communicate scientific knowledge.

Year 8 Science: Performance at the Low International Benchmark

Students performing at this level recognised some basic facts from the life and physical sciences. The multiple-choice item shown here illustrates the type of item likely to be answered correctly by students reaching the low benchmark. Students were given a definition of work (work is done when an object is moved in the direction of an applied force) and were asked to identify a diagram depicting a person doing work. On average internationally, this item was answered correctly by 78 per cent of students, who recognized that a person pushing a cart up a ramp was doing work. Every country except Tunisia had more than half their students answer correctly. Eighty-six per cent of Australian students answered this item correctly, significantly higher than the international average.



Country	Per cent full correct
Singapore	96 (0.9)
United States	91 (1.0)
Bulgaria	91 (2.1)
Russian Federation	91 (1.3)
Australia	86 (1.6)
England	85 (1.7)
International average	**78 (0.3)**
Tunisia	49 (2.1)

Year 8 Science: Performance at the Intermediate International Benchmark

Students reaching this benchmark were able to recognise and communicate basic scientific knowledge across a range of topics.

For this example, from the *biology* domain, the multiple-choice item required students to identify an animal characteristic found only in mammals. On average internationally, 63 per cent of Year 8 students recognised glands that make milk as the correct answer. Students in Australia, England and the United States performed poorly on this item, with the average for each of these three countries significantly lower than the international average. Particularly for Australian students, living in a country full of mammals, this result is surprising.

Which characteristic is found ONLY in mammals?

Ⓐ eyes that detect colour

● glands that make milk

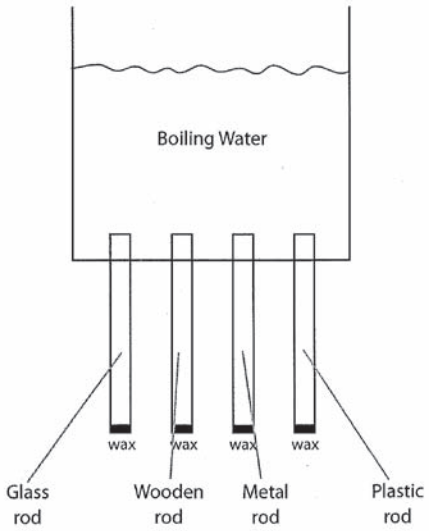
Ⓒ skin that absorbs oxygen

Ⓓ bodies that are protected by scales

Country	Per cent full correct
Chinese Taipei	91 (1.3)
Hong Kong	86 (1.8)
Thailand	84 (1.5)
International average	**63 (0.3)**
Australia	56 (2.7)
United States	53 (1.8)
England	53 (2.4)
Ghana	31 (2.1)

Year 8 Science: Performance at the High International Benchmark

Students reaching this benchmark were able to recognise and communicate basic scientific knowledge across a range of topics. This example shows the type of *physics* item likely to be answered correctly by students reaching the high benchmark. In the context of an investigation into thermal conductivity, this multiple-choice question asked students to choose among glass, wood, metal, and plastic for the best conductor of heat. On average, internationally, 47 per cent of students correctly chose metal as the best conductor. More than 70 per cent of students answered correctly in Singapore (79%) and in Chinese Taipei (75%). Students in Australia (60%), England (66%) and the United States (57%) performed on average significantly higher than the international average.



The diagram shows four identical size rods each of a different material sealed into the bottom of a container. The same amount of wax is placed on the end of each rod and then the container is filled with boiling water. On which rod will the wax melt first?

Ⓐ Glass rod
Ⓑ Wooden rod
● Metal rod
Ⓓ Plastic rod

Country	Per cent full correct
Singapore	79 (1.7)
Chinese Taipei	75 (1.8)
Japan	68 (1.9)
England	66 (2.3)
Australia	60 (2.5)
United States	57 (1.8)
International average	**47 (0.3)**
Indonesia	21 (2.1)

Year 8 Science: Performance at the Advanced International Benchmark

Students achieving at or above the advanced benchmark demonstrated a grasp of some complex and abstract science concepts. For this example, students were told that two substances together had a mass of 110 grams, and were asked to predict the mass of a new substance formed by combining the two original substances and explain their reasoning.

On average across countries only 23 per cent of students obtained full credit for their answer. In Japan, Korea and Chinese Taipei, more than half of the students answered correctly, while in Australia the proportion answering correctly was 25 per cent, not significantly different to the international average.

The mass of substances A and B are measured on a balance, as shown in Figure 1. Substance B is put into the beaker and substance C is formed. The empty beaker is put back on the balance, as shown in Figure 2.

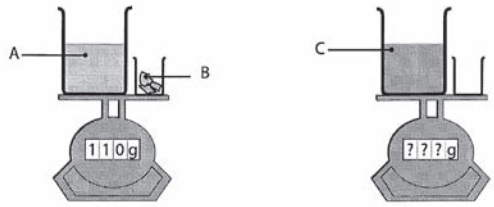


Figure 1 Figure 2

The scale in Figure 1 shows a mass of 110 grams.

What will it show in Figure 2?

(Tick one box.)

- ☐ More than 110 grams
- ☑ 110 grams
- ☐ Less than 110 grams

Explain your answer.

The two substances are still on the scales, one has just dissolved in the other. As both substances are still there it will still weigh the same.

Country	Per cent full correct
Japan	65 (2.1)
Korea	51 (2.0)
Chinese Taipei	51 (2.3)
England	28 (2.1)
Australia	25 (2.4)
United States	24 (1.6)
International average	**23 (0.3)**
Botswana	1 (0.4)

International results on the TIMSS science assessment

To place students' responses in a wider context, the item breakdown for Australian students was compared with the responses from students in other countries. Two countries were chosen for this international comparison. The first was Chinese-Taipei. Chinese-Taipei consistently performed in the top three of the 36 countries that participated at Year 4, and the 49 countries that participated at Year 8. Comparison with these students' responses provided an 'upper benchmark' for Australian students. The second country chosen was the United States. The US and Australia are often compared to one another because of curriculum and general cultural similarities.

Figure 3 shows TIMSS 2007 average levels of achievement for science in Years 4 and 8 for Australia, Chinese-Taipei and the US. In both figures, the TIMSS scale average is 500. The box below shows how to read the graphs.

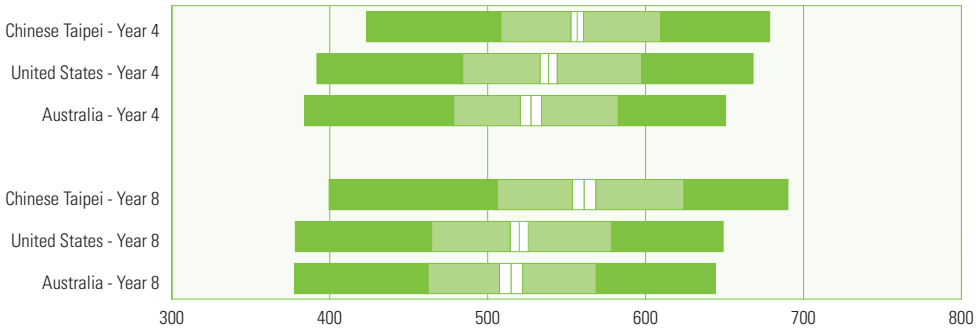
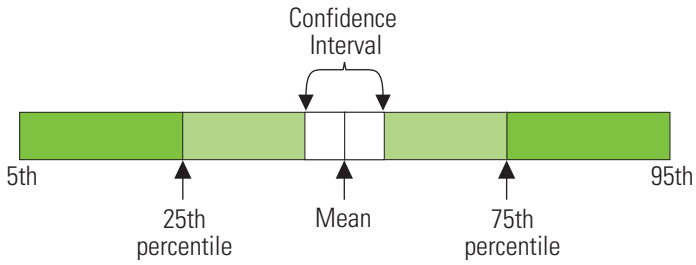


Figure 3 TIMSS 2007 achievement in Year 4 and Year 8 science

Figure 3 shows that for both Year 4 and Year 8 the average achievement level of Australian students was significantly higher than the TIMSS scale average. Australia's Year 4 performance was similar to that of eight other countries, significantly higher than 19 countries, but also lower than eight countries, including Chinese-Taipei and the US. Australia's average score for Year 8 was similar to three other countries, including the US, significantly higher than 35 countries and significantly lower than ten countries, including Chinese-Taipei.



READING THE GRAPHS

Each country's results are represented in horizontal bars with various colours. On the left end of the bar is the 5th percentile – this is the score below which 5 per cent of the students have scored. The next line indicates the 25th percentile. The next line at the left of the white band is the lower limit of the confidence interval for the mean – i.e. there is 95 per cent confidence that the mean will lie within this white band. The line in the centre of the white band is the mean. The lines to the right of the white band indicate the 75th and 95th percentiles.

Figure 4 shows the percentage of students from Australia, Chinese-Taipei and the US who reached the four benchmark levels in science.

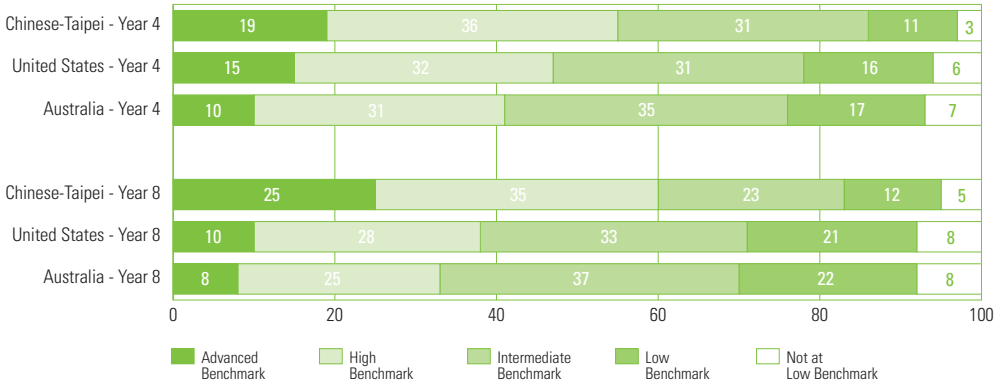


Figure 4 Percentage of students reaching the international benchmarks in Year 4 and Year 8 science

Figure 4 shows that Chinese-Taipei had the highest percentage of students achieving at the advanced benchmark, and the lowest percentage of students not reaching the low benchmark, for both Year 4 and Year 8 science. Australia and the US shared similar or the same rates of students not reaching the low benchmark but differed in terms of attainment at other levels. Australia had the lowest percentage of students achieving at the advanced benchmark for Year 4 and Year 8 science. Forty-one per cent of Year 4 students and 33 per cent of Year 8 students reached the high international benchmark.

In addition to average achievement levels, benchmark status was also related to each achievement item contained in the TIMSS 2007 assessment. In other words, every item was linked to attainment of one of the four benchmarks. The associated benchmark for each item discussed in this report is listed during the item analyses.

How can TIMSS results inform pedagogy for classroom teachers?

The main purpose of this report is to present TIMSS 2007 science results in a way that can inform pedagogy. The report explores students' responses to a selection of science items and then considers what these responses might indicate about students' level of understanding for a particular item and its content area. By breaking down the results at the item level, teachers can ascertain whether the mistakes typically made by students in the sample are also mistakes made by their own students.

Not all items included in the TIMSS 2007 project are available to the public. The selection of items discussed in this report was made from the group which have been publicly released. The CD included with this report contains all of the TIMSS 2007 released science items so that teachers may see the types of questions students completed when they participated in the project.

Using students' answers as evidence of their understanding

Item analysis of the TIMSS 2007 results can inform pedagogy because we assume that students' answers can be used as evidence of their understanding of scientific concepts. For instance, items presented via a multiple choice format offer a good opportunity to evaluate students' understanding in large scale studies like TIMSS. The series of options provided in a multiple-choice question contain the correct answer in addition to a set of distracter options. Some of the distracters represent answers that students will obtain if they have a basic understanding of the area, some if students have an intermediate understanding of the material and some are extreme answers that are likely picked if students guess. Therefore, if a majority of students select a distracter as the answer for an item, it is possible to discuss the average level of understanding at which Australian Year 4/Year 8 students are operating. This discussion is possible due to the fact that TIMSS is a large scale study that assesses the achievement of a representative student sample.

Another important factor to consider is the percentage of students that omit giving an answer for an item. There are two logical explanations for this. The first is that the information is missing at random; in other words, students missed the item by accident. The second explanation is that the information is missing for a reason. In this case, educators must ask – did students omit the answer because (i) they believed they did not have ability to obtain a solution or (ii) they did not want to put in the effort required to complete the task? This report will assume that it is the former, perceptions of doubt in relation to ability, that explains the percentage of omitted answers for an item, rather than a lack of effort, or because the information was missing at random.

Types of items and item analysis

Every item included in the TIMSS 2007 project had several assessment characteristics. Each corresponded to a curriculum area (*mathematics* or *science*), a content area (e.g. *algebra* or *chemistry*) and a cognitive skill (*knowing*, *applying* or *reasoning*)[1]. The mode of item presentation also varied. Items were either *multiple-choice* questions or questions requiring a *constructed–response*. The majority of items discussed in this report were of the multiple-choice format.

The following sections of this report present the item analyses for a selection of TIMSS 2007 results. Each item is presented along with a set of percentages. These percentages might include the number of students who obtained the correct/incorrect solution, the number of students who selected a particular distracter option, the number of students who omitted giving the answer to an item, or the percentage of gender difference in the number of correct answers given (e.g. if girls performed better than boys on average, by how large a percentage?). Rather than providing all these figures for every item, a set of percentages were selected for each that reflected the most

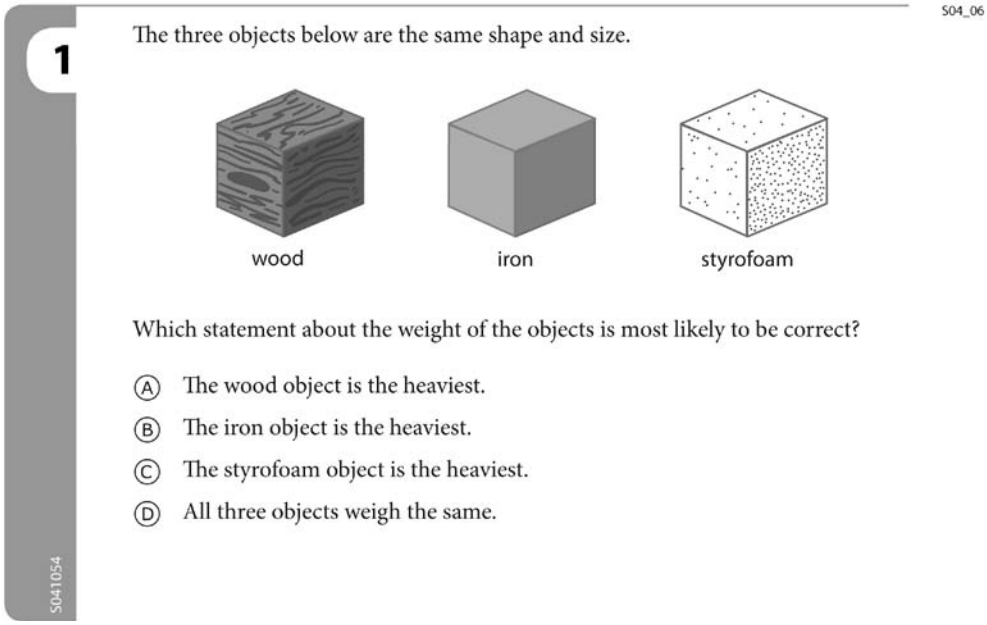
1 Note that items assessing the last of the cognitive domains, reasoning, are not considered in this report.

significant trend for that item. Note that all percentages discussed have been rounded to the nearest whole per cent except in the case of figures below 0.5%.

Part of the TIMSS 2007 project also included surveying the teachers of the participating students. Part of this survey asked these teachers to rate whether topics assessed in the TIMSS items had been (i) "mostly taught before this year"; (ii) "mostly taught this year"; or (ii) "not yet taught or introduced". For each item analysed in this report, the percentage of students who had teachers that rated 'yes' to the first two categories is given. Thus, we report the proportion of students that, according to their teachers, had been exposed to the relevant topics prior to their participation in TIMSS.

What can TIMSS tell us about Year 4 science?

Five science items from the set presented to Year 4 students are discussed in this report. The first of these was a *low benchmark* item. It investigated students' understanding of *physical science* in the *applying* domain. Students were presented with three objects of the same size and shape that varied in terms of the material they were made from – wood, iron or styrofoam.


S04_06

1

The three objects below are the same shape and size.

wood iron styrofoam

Which statement about the weight of the objects is most likely to be correct?

(A) The wood object is the heaviest.
(B) The iron object is the heaviest.
(C) The styrofoam object is the heaviest.
(D) All three objects weigh the same.

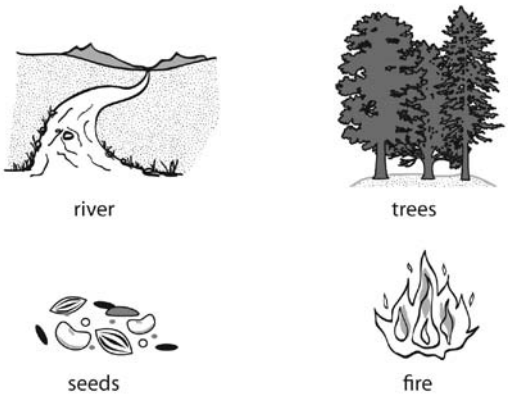
S041054

Ninety-one per cent of Chinese-Taipei students and 80% of US students were able to correctly identify that the iron object would be the heaviest. Only 68% of Australian students selected iron; however, just 44% of students had teachers who stated that the topic had been covered. Item 1 was also one which demonstrated a large gender difference in the percentage of correct responses for Australian students. Boys had a 19% higher correct response rate than girls. In the US, this trend in favour of boys was 7% and in Chinese-Taipei the number of girls who answered correctly was higher but only by 0.3%.

Items 2 and 3 assessed the *life science* domain. Item 2, associated with the *intermediate benchmark*, was an opportunity to investigate students' application *(applying)* of their knowledge in this area.

S02_01

2 Look at the four pictures.



Which pictures show **non-living** things?

Ⓐ trees and fire
Ⓑ fire and river
Ⓒ river and seeds
Ⓓ seeds and trees

S041007

Eighty-six per cent and 73% of Chinese-Taipei and US students, respectively, answered this item correctly compared to only 67% of Australian students. The teachers of 77% of Australian students reported that the classification of living things has been previously taught. For Australia (16%) and the US (13%), the most commonly chosen distracter was option C, *rivers and seeds*, indicating that some students perceived seeds to be non-living and/or could not differentiate between a river and the things that live in a river. Furthermore, by omission, these students believed fire was a living organism. Item 2 was also another to be associated with salient gender differences. Interestingly, these differences varied across countries. Girls in Australia and Chinese-Taipei had a higher correct response rate for the item by 7% and 6%, respectively. In contrast for the US, boys had an 8% higher correct response rate.

Item 3 examined students' knowledge *(knowing)* of *life science* with positive results for Australian students. This item was linked to the *low benchmark*.

S02_05

3 Smoking harms the body in many ways.

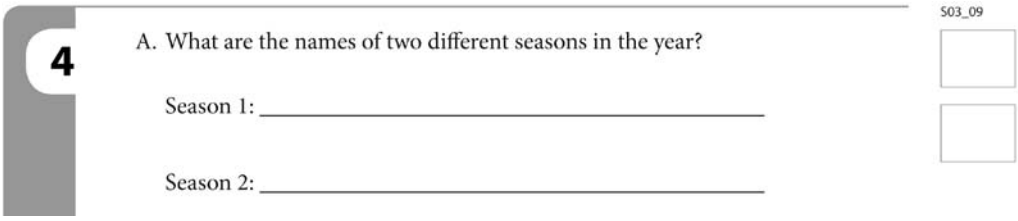
It is most harmful to which organ?

Ⓐ lung
Ⓑ kidney
Ⓒ liver
Ⓓ stomach

S041042

Eighty-nine per cent of US students and 87% of Australian students correctly identified that smoking is most harmful to the lungs; however, only 69% of Chinese-Taipei students made this connection. The increased awareness of US and Australian students may be a reflection of exposure to public media campaigns in addition to science curriculum. Eighty-eight per cent of Australian students had teachers who stated that ways of maintaining good health had been discussed in the classroom.

The final two Year 4 science items presented were from the *earth science* domain. In particular, they revolved around weather conditions, which the teachers of 69% of Australian students reported having covered. The first item, Item 4, investigated students' knowledge (*knowing*) of seasons and represented the *low benchmark*. Results for this item implied that the majority of students from all three countries had a good grasp of this concept.


4

A. What are the names of two different seasons in the year?

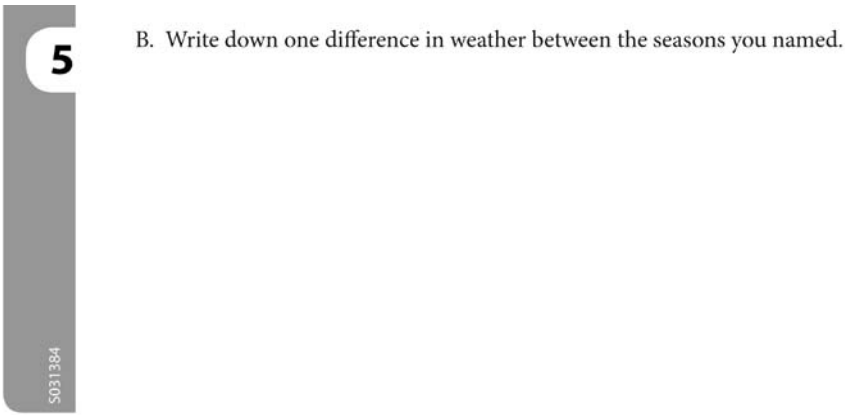
Season 1: _____

Season 2: _____

S03_09

Eighty-nine per cent of Australian students, 90% of US students and 93% of Chinese-Taipei students were able to name two seasons.

Item 5 was presented directly after Item 4 and designed as a follow-up assessment of students' understanding of seasons. It also accessed more higher-order cognitive processes (*applying*) and was linked to the *intermediate benchmark*.


5

B. Write down one difference in weather between the seasons you named.

S031384

Responses to Item 5 illustrated the limitations of students' understanding when they were asked to apply this knowledge. Only 68% of Australian students and 69% of US students could describe a difference between the two seasons they had nominated. This number was much higher for Chinese-Taipei students (81%). Whereas 5% of Australian students had avoided naming two seasons for Item 4, 11% of Australian students omitted giving an answer to distinguish between them for Item 5. Across all three countries, there were higher numbers of girls who answered Item 5 correctly – 5% for Chinese-Taipei, 10% for the US and 18% for Australia.

What can TIMSS tell us about Year 8 science?

The first two of the five Year 8 science items assessed students understanding of *biology* in the *knowing* cognitive domain. Item 6, an item associated with *high benchmark*, examined students' knowledge of genetics; however, only 28% of Australian students had teachers who reported covering the topic of heredity and reproduction.

S05_10

6 Which of the following is the best way to determine whether two people are related?

Ⓐ Compare their blood types.
Ⓑ Compare their handwriting.
Ⓒ Compare their genes.
Ⓓ Compare their fingerprints.

S032035

Considering the small proportion of students exposed to the topic, Australian students' performed relatively well with 71% of students selecting the correct answer. This was compared with 86% of Chinese-Taipei students and 67% of US students. Twenty-four per cent of Australian students and 27% of US students chose comparing blood types as the appropriate answer. Another interesting trend for this item was related to gender differences. The correct response rate for Australian boys was 8% higher than that for Australian girls. The correct response rate for boys in Chinese-Taipei was only 0.2% higher and in the US, the rate was 3% higher for girls.

Item 7 investigated students' understanding of reproductive *biology*. This item represented the *advanced benchmark*.

S02_05

7 Which of the following is formed immediately after fertilisation?

Ⓐ egg
Ⓑ sperm
Ⓒ zygote
Ⓓ embryo

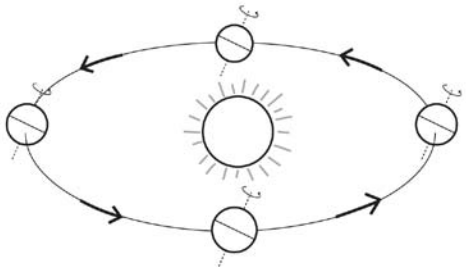
S042028

Only 15% of Australian students selected *zygote*, compared with 25% of US students and 81% of students in Chinese-Taipei. The incorrect response rate for this item was 84% for Australian students, 73% for US students and only 15% for students in Chinese-Taipei. More than half of Australian and US students who answered incorrectly for this item (44% of Australian students and 42% of US students) believed that an embryo is formed immediately after fertilization.

The third Year 8 science item, Item 8, examined the *applying* cognitive skill and was from the area of *earth science*. Like Item 7, this question was from the pool of *advanced benchmark* items and drew on knowledge relating to earth's rotation around the sun. Fifty-seven per cent of Australian students had teachers who stated that the movement of bodies in the solar system, and related phenomena, was a topic that had been discussed before participation in the TIMSS 2007 project.

S07_09

8



The diagram above shows the Earth's path around the Sun and the tilt of Earth's axis. Which of the following patterns on Earth is caused by the tilt of Earth's axis?

(A) seasons
(B) day and night
(C) years
(D) time zones

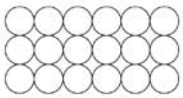
S032663

Fifty-one per cent of Australian students were able to identify that seasonal changes were due to the tilt of the earth's axis as it rotates around the sun. Sixty-seven per cent of US students also recognised this fact relative to only 24% of Chinese-Taipei students. Fifty-one per cent of Chinese-Taipei students selected *'day and night'* compared with 21% of US students and 27% of Australian students.

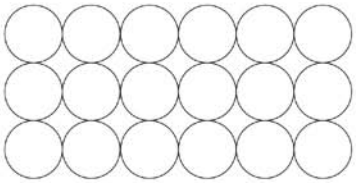
Item 9 was selected from the *physics* domain and also assessed students' *application* ability in this content area. This item was particularly challenging and, accordingly, selection of the correct solution classified a student's response as above the *advanced benchmark*. The teachers of 50% of Australian students reported that temperatures changes related to changes in movement or speed of particles had been discussed in the classroom.

S04_08

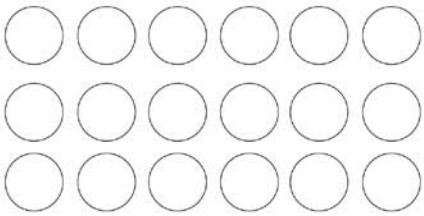
9 The diagram represents the arrangement of particles in a metal before it has been heated.



Which diagram represents the arrangement of particles in the metal after it has been heated?

Ⓐ 

Ⓑ 

Ⓒ 

Ⓓ 

S042061

Australian students' had the best performance results for this item, but not surprisingly, the general correct response rates were quite low. Forty-four per cent of Australian students, 40% of US students and 27% of Chinese-Taipei students correctly identified option D as the answer and were able to apply their knowledge to see that the particles would become less packed after heating. Fifty-five per cent of Australian students, 60% of US students and 73% of Chinese-Taipei answered incorrectly for this item.

The last Year 8 science item, Item 10, examined *chemistry* in the *knowing* domain. Item 10 represented the *advanced benchmark*.

S05_05

10 Which is NOT an example of a chemical change?

Ⓐ Melting ice

Ⓑ Corroding silver

Ⓒ Burning match

Ⓓ Rotting vegetation

S022208

Thirty-seven per cent of Australian students correctly selected option A along with 47% of US students and 74% of Chinese-Taipei students. Sixty-two per cent of Australian students, 52% of US students and 26% of Chinese-Taipei students answered incorrectly, and therefore, by omission, believed that melting ice was an example of a chemical change in action. This item was also associated with large gender differences but only for Australian students. Australian boys' correct response rate for Item 10 was 9% higher than the rate for girls. Boys in the US and Chinese-Taipei had correct response rates 0.8% and 2% higher, respectively. Fifty-nine per cent of Australian students had teachers who stated the topic of chemical changes had been discussed prior to students' participation in the TIMSS 2007 project.

Summary and general trends

The items discussed in this report represent a small selection of the total item inventory that was part of the TIMSS 2007 project; however, investigation of these items illustrated areas of strength and (particularly) weakness for Australian students that warrant educators' consideration. While it is difficult to identify trends based on a small sample of items, there were also some central themes that emerged through the item analysis conducted.

At the national level

The Year 4 science items represented a selection from the areas of physical science, earth science and life science. Australian students' performance in the latter two areas was variable, with deficiencies revealed when an item required the application of knowledge. This result suggests that the quality of students' knowledge may be lacking; students were able to answer questions that tapped into the surface level of a content area but struggled when items required a deeper understanding.

For Year 8 science, biology, earth science, physics and chemistry were examined in the five items presented. Again Australian students' performance was variable with high levels of incorrect responses relating to items examining reproductive biology and chemistry.

For some of the items discussed the percentage of answers omitted was quite large. Avoidance of these items is an issue of concern whether it was due to poor competence beliefs or lack of effort. Lastly, larger gender differences for the rate of correct responses tended to favour boys. This trend meets with the general TIMSS 2007 finding that boys outperformed girls in science and mathematics.

At the international level

For the most part, the international comparisons made between Australia, the US and Chinese-Taipei served to highlight areas of weakness for Australian students. There was an obvious gap between the understanding of Australian students and their Chinese-Taipei counterparts. Furthermore, where Australian and US students previously achieved at similar levels (see TIMSS 2003 results), in TIMSS 2007 the US outperformed Australia in almost all the items reviewed.

On the other hand, the international comparison made also illustrated the trap of simply considering average level results for a country. In three of the ten items reviewed, Chinese-Taipei was outperformed by either the US or Australia. This finding does not diminish the high achievement levels of the Chinese-Taipei students who participated in TIMSS 2007 but emphasizes that, while they were ranked in the top three for all grade/curriculum areas, Chinese-Taipei students still demonstrated skill deficiencies in some areas.

Informing pedagogy

The primary aim of this report was to provide an informative review of Australian students' performance in TIMSS 2007. This analysis was conducted with the hope that teachers might reflect on the results and that this might be helpful to their classroom teaching initiatives.

Martinez[2] (2001) pointed out that it is not useful to consider TIMSS as an international achievement competition. Rather it is a "compendium of curricular data, educational cultures, teaching and learning styles, and assessment techniques" (p.114). The TIMSS 1999 Video Study of Eighth Grade Mathematics Teaching demonstrated both the similarities and discrepancies in teaching amongst eight participating countries that included Australia and the US. For instance, Australia and Japan spent more time practicing new content in the classroom than the US who devoted more time to reviewing material. Thus, the strengths and weaknesses of students' responses highlighted in this report were likely due to many different factors. However, with more awareness of students' understanding in different curriculum areas, educators can develop learning strategies that suit their particular teaching styles and unique educational contexts.

More information!

Included with this report is a CD which contains all of the released science items for Year 4 and Year 8, along with the scoring guide for each item. Teachers can use these to see how items on international assessments are constructed and scored, and use the formats for their own testing.

Also on the CD are the item almanacs for all of the released items. These show, for each item, the number of students who attempted the item, the percentage of students who responded to each of the various marking codes, the percentage of students who omitted the item and the percentage of students who did not reach it. Also included are the percentage of students overall who gave a correct answer, and the percentage of boys and girls in each country giving a correct answer. We encourage teachers to explore these statistics for themselves.

Further information and all reports on all TIMSS assessments is available from the TIMSS website, at www.acer.edu.au/timss.

2 Martinez, J. (2001) Exploring, inventing, and discovering mathematics: A pedagogical response to TIMSS. *Mathematics Teaching in Middle School, 7 (2)*, 114-120.

www.acer.edu.au