Overview of the Third International Mathematics and Science Study (TIMSS) in Australia

Jan Lokan

SCOPE OF TIMSS

TIMSS is the largest, most comprehensive study of educational achievement ever undertaken. More than half a million students from over 15 000 schools in 45 countries participated in the study, as did their mathematics and science teachers and the principals of their schools.

The students were sampled randomly from three target populations in each country. The target populations were defined as:

Population 1: the two adjacent grade levels containing the largest proportion of nine-year-old students at the time of testing

Population 2: the two adjacent grade levels containing the largest proportion of thirteen-year-old students at the time of testing

Population 3: the final year of secondary schooling.

In Australia, these definitions meant that in some states the Population 1 students came from Years 3 and 4 (in NSW, Victoria, Tasmania and the ACT) and in some they came from Years 4 and 5 (in Queensland, SA, WA and the NT). Similarly, the Population 2 students came from Years 7 and 8 or Years 8 and 9. The Population 3 students were in Year 12 in all states and territories. Over 29 000 Australian students, from about 1500 classes in almost 450 schools, participated in TIMSS. The Population 1 and Population 2 testing took place in Australia late in 1994 and the Population 3 testing was done late in 1995. In northern hemisphere countries all three populations were tested from March to May 1995 (almost at the end of their school year, as was also the case in Australia).

WHY DO INTERNATIONAL STUDIES?

The idea behind international studies such as TIMSS is that they make use of naturally existing differences which would not be possible to implement within a single country. Countries differ in the ways their school education is organised, in the curricula they offer, in the preparation required of their teachers, in the styles their teachers use to present the curricula, in their expectations of students, and in many other factors potentially related to effective teaching and learning. In the final secondary years, curriculum differences and differences in the organisation of schooling are even more pronounced than they are at earlier stages (for example, in many European countries the senior secondary students attend schools with very different types of programs).

Well designed international studies can provide information on what is possible for students to achieve and what conditions are most likely to facilitate their learning. Whatever a country's reasons for taking part in a study like TIMSS, the underlying challenge will always be to determine more about effective school organisation and effective teaching and learning.

WHAT DATA WERE COLLECTED IN TIMSS?

Altogether, at Population 1, there were 107 mathematics test questions and 101 science test questions; at Population 2 there were 157 mathematics and 140 science questions; and at Population 3 there were 83 'mathematics and science literacy', 68 advanced mathematics and 67 physics questions. (The Population 3 results had not been released at the time of the ACER conference, and hence the remainder of this paper is concerned with Populations 1 and 2 only.)

The test questions were divided up into several booklets, which were allocated to the students at random, so that equivalent groups of students answered each test booklet and each student answered only one booklet. The majority of the test questions were multiple choice format, but about a third of the testing time was taken up with questions to which the students had to construct their answers.

In addition, there were questionnaires for students, their mathematics and their science teachers, and the principals of their schools. In these 'context' questionnaires, more than 1500 questions were asked. Each question related to an aspect of education that was thought to be important in relation to student achievement, often because it had been found to be so in previous research.

The TIMSS tests and questionnaires were translated into 31 languages, for use by the 45 participating countries.

WHAT DID THE TESTS ASSESS?

Items in the TIMSS tests assessed the objectives and skills that were thought to be important on the basis of curriculum analyses undertaken in most of the countries which participated in the Population 2 testing.

Content areas in mathematics were:

- Fractions and number sense;
- Algebra (Patterns, relations and functions at Population 1);
- Geometry;
- Data representation, analysis and probability;
- Measurement; and
- Proportionality (Population 2 only).

Content areas in science were:

- Life science;
- Physics;
- Earth science;
- Chemistry; and
- Environment and Nature of science.

Most of the skills assessed were similar in mathematics and science:

- Knowing;
- Using routine procedures;
- Performing complex procedures;
- Understanding simple information;
- Understanding complex information;
- Theorising, analysing and solving problems; and
- Investigating the natural world (science only).

Some of the investigative aspects were assessed more deeply in the 'performance assessment' component of TIMSS which was carried out in a much smaller sample of schools and students (a subsample of participants) than the main study.

WHAT WERE THE MAIN RESEARCH QUESTIONS THAT TIMSS WAS SEEKING TO ANSWER?

In keeping with TIMSS's conceptualisation of the curriculum as occurring at three levels, three of the four general research questions that guided the development of the study focused on the curriculum levels:

1 The intended curriculum

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?

2 The implemented curriculum

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

3 The attained curriculum

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?

The fourth general research question incorporated all curriculum levels in relation to the contexts in which schooling occurs:

4 Relationships of curricula to social and educational contexts

How are the intended, implemented, and attained curricula related with respect to the contexts of education, the arrangements for teaching and learning, and the outcomes of the educational process?

WHAT WERE AUSTRALIA'S RESULTS?

It is possible here to describe only a few features of Australia's results, which are reported in detail in two books of more than 200 pages each.[i]

Relative standings overall

In terms of relative standing, our Australian students performed well, equal to or better than their peers in other English speaking countries and in many European and other countries. Relatively, our standing in science was rather better than our standing in mathematics, and our primary students (Population 1) performed slightly better than our lower secondary students (Population 2). The main area of concern for us is that, in both populations, the students from our neighbouring Asian countries of Singapore, Korea, Japan and Hong Kong outperformed our own students in mathematics, by quite a large margin, and students from the first three of these countries usually outperformed our students in science. Countries performing significantly better than Australia are shown by population in Table 1.

Table 1 Countries Performing Better than Australia

MATHEMATICS		SCIENCE	
Population 2 upper grade			
Singapore	Belgium (Flemish)	Singapore	Japan
Korea	Czech Republic	Czech Republic	Korea
Japan	Slovak Republic		
Hong Kong	Switzerland		
Population 2 lower grade			
Singapore	Belgium (Flemish)	Singapore	Bulgaria
Korea	Czech Republic	Korea	Slovenia
Japan	Netherlands	Czech Republic	Belgium (Flemish)
Hong Kong		Japan	
Population 1 upper grade			
Singapore	Hong Kong	Korea	Japan
Korea	Netherlands		
Japan	Czech Republic		
Population 1 lower grade			
Korea	Japan	Korea	
Singapore	Hong Kong		

Highest achieving students

An alternative perspective on relative performance is provided in Figures 1 and 2. These show, for a subset of the participating countries, how each country contributed to the highest scoring ten per cent of students from all countries combined. For example, of the highest scoring ten per cent in Population 2 mathematics around the world (about 27 000 students), over 3700 were from Singapore, which was about 45 per cent of the participating students from that country. The highest scoring ten per cent worldwide contained about 13 per cent of the Australian students in Population 2 mathematics, about 18 per cent in Population 2 science, about 14 per cent in Population 1 mathematics and about 16 per cent in Population 1 science. So we can say that we were ‘a little ahead of the game’, so to speak, but definitely not what might be regarded as ‘outstanding’.

Perhaps we are not doing as much for our best students as we could, to challenge them and help them to achieve their potential.

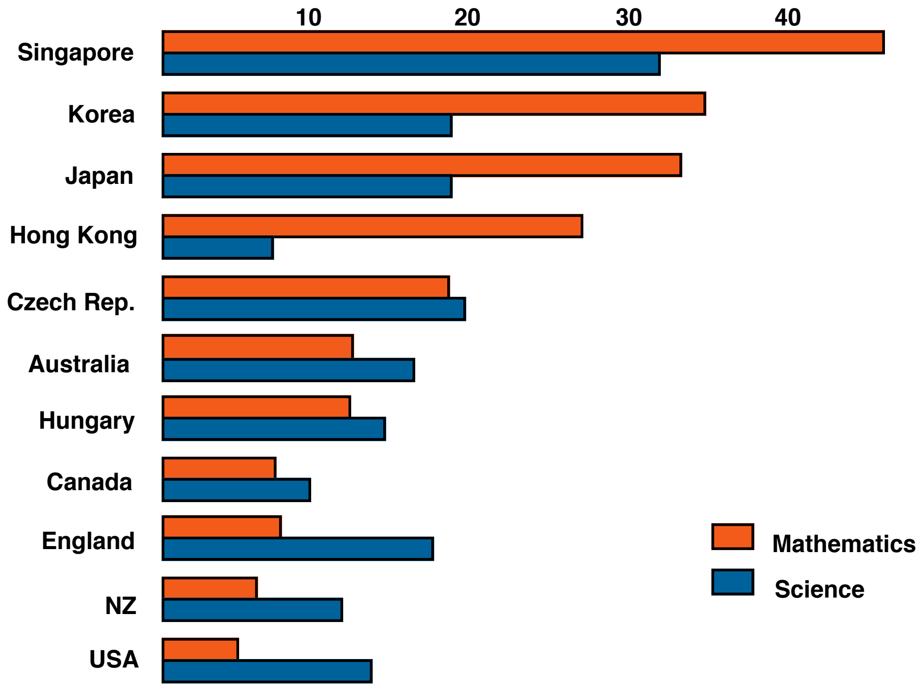


Figure 1 Country Percentages of Students in World's Top Ten Per Cent, Population 2

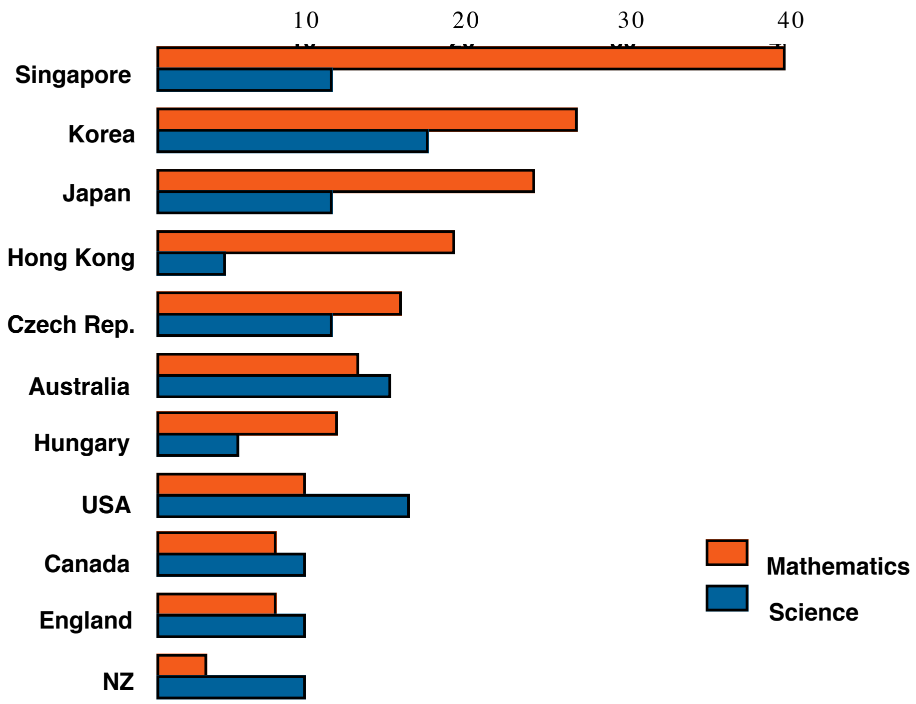


Figure 2 Country Percentages of Students in World's Top Ten Per Cent, Population 1

Figures 1 and 2 are also interesting in that they highlight the greater spread of 'top ten per cent' students worldwide among countries in mathematics than in science. They also show how much better some countries did in mathematics than in science (particularly the Asian countries) together with the converse of countries doing somewhat better in science than in mathematics (particularly the USA and New Zealand in both populations and England in Population 1).

Content area results

Some interesting comments can also be made about Australia's performance in the various content areas represented in the TIMSS tests. Our students achieved above average or average results in all areas tested, at both populations. At Population 2, our achievement was furthest above average (relative to ourselves) in the mathematics areas of 'algebra' and 'chance and data' and in the science areas of 'environmental issues', 'nature of science' and 'physics'. At Population 1 we were furthest above average in the mathematics areas of 'measurement' and 'geometry' and the science areas of 'environmental issues' and 'nature of science'.

Results by gender

Australia shared the distinction with only a handful of other countries of having our boys and girls performing at equivalent levels in both mathematics and science. Of eight main gender comparisons, by upper and lower grade within population on each of the mathematics and science tests, only one gender difference occurred in Australia—at Population 1, our upper grade boys achieved significantly higher science scores than the upper grade girls. There were only four countries where no gender difference was found in any of these eight possible comparisons—Cyprus, Ireland, Singapore and Thailand. Internationally, gender differences were not common in mathematics but were pervasive in science. Australia can take heart that our efforts over the past 10 years or so to make mathematics and science instruction more gender equitable seem, at least from the TIMSS results, to have largely achieved that purpose.

State and territory results

Just as our results were at or above the international average in both mathematics and science overall, the same was true for our results considered separately by state. There was considerable spread between our highest and lowest performing states, but even the lowest achieved results at the international average. Our highest achieving states performed at the level of the highest achieving countries worldwide, with WA sharing top position with Korea in Population 1 science.

We all recognise that it is difficult to make comparisons of results among the Australian states and territories because there are many contextual factors that need to be taken into account. The same is true internationally, of course. Internationally, TIMSS took the view that actual results should be reported, but they should be accompanied by a wide range of contextual data to allow countries to interpret results in as informed a way as possible. The same practice was adopted in reporting results for the Australian states and territories, each of which is best informed about the contextual factors underlying its own education system and policies — differences in school starting age, for example.

I will return later to a discussion of relationships between achievement and some of the contextual factors measured in the TIMSS questionnaires. At this stage, while we are thinking about Australia's results, it will be useful to focus on some examples of individual TIMSS test questions and on how the Australian students performed on these. I have chosen questions which illustrate the full range of our students' performance, from some on which we were best in the world, some on which our performance was average internationally, and some on which our performance was lowest in the world. Where it is of particular interest, I have shown results for Singapore (the highest achieving country in all areas but Population 1 science) or for the four Asian countries combined.

There are messages for our policy makers in some of the examples and results shown. The items on this page are all ones on which Australia scored substantially below the four Asian countries, even if close to the international average on some. These items, plus those on the next three pages, are from the *Population 1* test.

Subtract: $\begin{array}{r} 6000 \\ -\ 2369 \\ \hline \end{array}$

A 4369 **C** 3631

B 3742 **D** 3531

Aust: 47 %
Int'l: 71 %
Asian: >90 %

- -

Add: $\begin{array}{r} 6971 \\ +\ 5291 \\ \hline \end{array}$

A 11 162 **C** 12 262

B 12 162 **D** 1 211 162

Aust: 76 %
Int'l: 84 %
Asian: 95 %

<u>Addition Fact</u>

4 + 4 + 4 + 4 + 4 = 20

Write this addition fact as a multiplication fact.

______ x ______ = ______

Aust: 71 %
Int'l: 77 %
Asian: >90 %

- -

25 x 18 is more than 24 x 18. How much more?

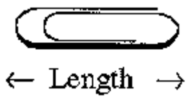
A 1 **C** 24

B 18 **D** 25

Aust: 41 %
Int'l: 45 %
Asian: >70 %

On the next two items, which teachers here would say are more closely related to our curriculum, Australia did score above the international average—at a similar level to the Asian countries on the first of the two but below the Asian countries on the second.

Here is a paper clip.

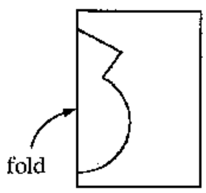


About how many lengths of the paper clip is the same as the length of this line?

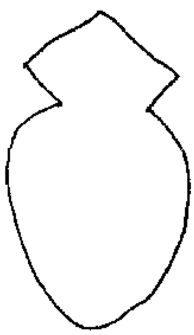
Answer: 4 paper clips

Aust: 58 %
Int'l: 48 %
Asian: 60 %

Craig folded a piece of paper in half and cut out a shape.



Draw a picture to show what the cut-out shape will look like when it is opened up and flattened out.



Aust: 72 %
Int'l: 59 %
Asian: 85 %

The first item below, a science item, is interesting because of the differences between the Asian countries, with Hong Kong and Japan performing much above the international average, Singapore (with Australia) at the international average and Korea considerably below this level. The second item is a further mathematics item on which Australia performed at the international average, but much below the Asian countries.

Which of these would most likely be measured in millilitres ?

A **the amount of liquid in a teaspoon**

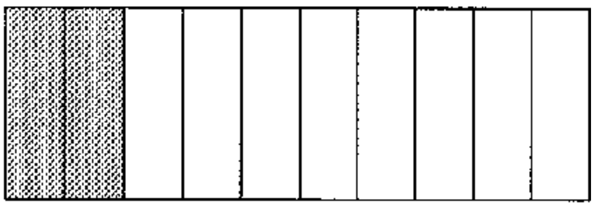
B **the weight of a pin**

C **the amount of petrol in a tank**

D **the thickness of 10 sheets of paper**

Aust: 45 %
Int'l: 38 %

HK/Japan: 75 %
S'pore: 45 %
Korea: 31 %



Which number represents the shaded part of the figure?

A. 2.8

B. 0.5

C. 0.2

D. 0.02

Aust: 40 %
Int'l: 40 %
Asian: >70 %

The three items illustrated on this page, all Population 1 science items, are ones on which the Australian students performed well, much above the international average. The items relate to life science or environmental science, areas of strength for us. The Asian countries had varying performances on these items, some above average and some well below. In the main, Japan and Korea performed better on these items than Singapore and Hong Kong.

The human brain is inside the skull. Write down one advantage of the skull being thick and strong.

The skull protects the brain so you won't have many bad injuries like a hand thrown at your head.

Aust: 66 %
Int'l: 51 %

Write down one thing your heart does that helps the other parts of your body.

Your heart pumps blood to the muscles and that helps you move.

Aust: 69 %
Int'l: 40 %

Write as much as you can about why large oil spills in rivers and seas are harmful to the environment.

Because it cause dangers to all sea and river life by killing them off then all the other animals die to from eating posioned food which starts a chain reaction. It also pollutes our waterway.

Aust: 75 %
Int'l: 58 %

Now we have moved to items from the Population 2 tests. On the kinds of items shown on this page, Australia performed either worst of all countries or close to worst. The Asian countries all performed very highly. Division of fractions, items involving more than two fractions and most items involving decimals were areas of weakness in Australia. A teacher commented to me recently that 'we would not expect students to do the decimal division without a calculator', but, if you know what you're doing, the item involves no more than division by 4—a skill which it could be argued that Year 8 and 9 students should have.

$$\frac{3}{4} + \frac{8}{3} + \frac{11}{8} =$$

A $\frac{22}{15}$

B $\frac{43}{24}$

C $\frac{91}{24}$

D $\frac{115}{24}$

Aust: 35 %
Int'l: 50 %

Divide: 0.004) 24.56

A 0.614

B 6.14

C 61.4

D 614

E 6140

Aust: 23 %
Int'l: 44 %

Divide: $\frac{8}{35} \div \frac{4}{15} =$

$$\frac{\cancel{8}^{2}}{35} \times \frac{15}{\cancel{4}_{1}} = \frac{2}{35} \times \frac{15}{1}$$

$$= \frac{30}{35}$$

$$= \frac{6}{7}$$

Aust.: 25 %
Int'l: 43 %

Performance on the three items shown on this page was generally quite low around the world. The Asian countries all scored about 80 per cent correct on the mathematics item shown below and on the science item about the splint bursting into flame. On the item about which gas is found in the greatest amount in air, the Asian countries were still substantially above average, each scoring between about 50 and 60 per cent correct.

Subtract: $\frac{2x}{9} - \frac{x}{9} =$

A $\frac{1}{9}$

B 2

C x

D $\frac{x}{9}$

E $\frac{x}{81}$

Australia: 48 % correct
International: 51 % correct

Air is made up of many gases.
Which gas is found in the greatest amount?

A nitrogen

B oxygen

C carbon dioxide

D hydrogen

Aust: 16 %
Int'l: 27 %

Which gas could cause a glowing splint to burst into flame?

A neon

B oxygen

C nitrogen

D carbon dioxide

Aust: 38 %
Int'l: 50 %

The science items on this page were done relatively well by the Australian students, in relation to the international average performance. The Asian countries performed at the same level as Australia, or slightly lower, on the first item shown here. On the item about the unwanted consequences of introducing a new species to an area, Australia performed highest in the world, considerably higher than the Asian countries (especially Hong Kong and Japan, which performed below the international average).

Animals are made up of many atoms.
What happens to the atoms after an animal has died?

A The atoms stop moving. (13 %)

B The atoms recycle back into the environment. (36 %)

C The atoms split into simpler parts and then combine to form other atoms. (8 %)

D The atoms no longer exist once the animal has decomposed. (43 %)

International average (upper grade): 26 % correct

What could be the unwanted consequences of introducing a new species to a certain area? Give an example.

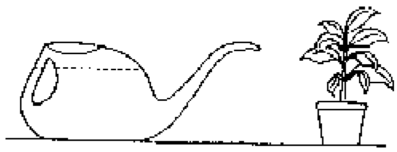
An ecological disaster.

Eg. - Rabitts eat food meant for native animals. Native animals starve and rabitts multiply so all natives die and lots of rabitts live.

Aust: 74 %
Int'l: 37 %

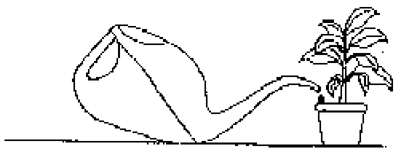
The novel science item featured on this page was one on which the Australian students' performance was at the international average. Singapore and Hong Kong students performed very well on this item; Japanese and Korean students performed at about the same level as Australia. The correct answer for this item is shown on the left hand side of the lower diagram. The most common wrong answer was to draw the line parallel to the bottom of the tilted can. Several of the more unusual wrong answers are illustrated on this page.

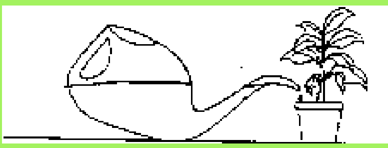
A watering can is partly filled with water as shown.



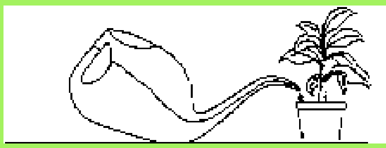
The watering can is tipped so that the water just begins to drip through the spout.

Draw a line to show where the surface of the water in the can is now.

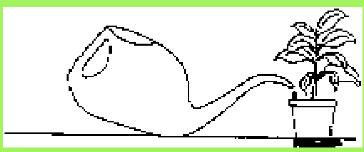




Per cent correct



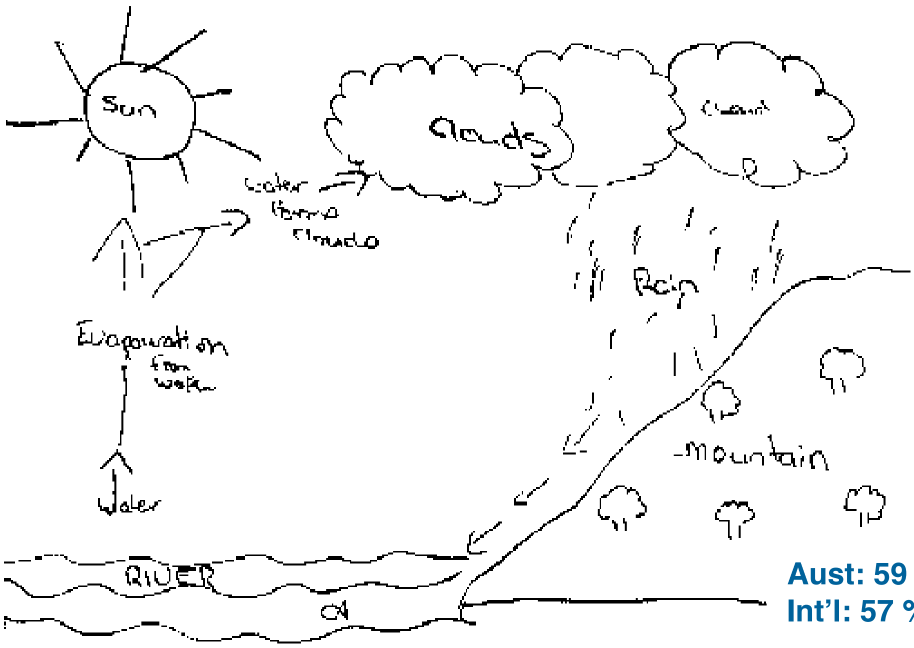
Pop. 2 Aust: 57
Int'l: 52



Pop. 1 Aust: 20
Int'l: 21

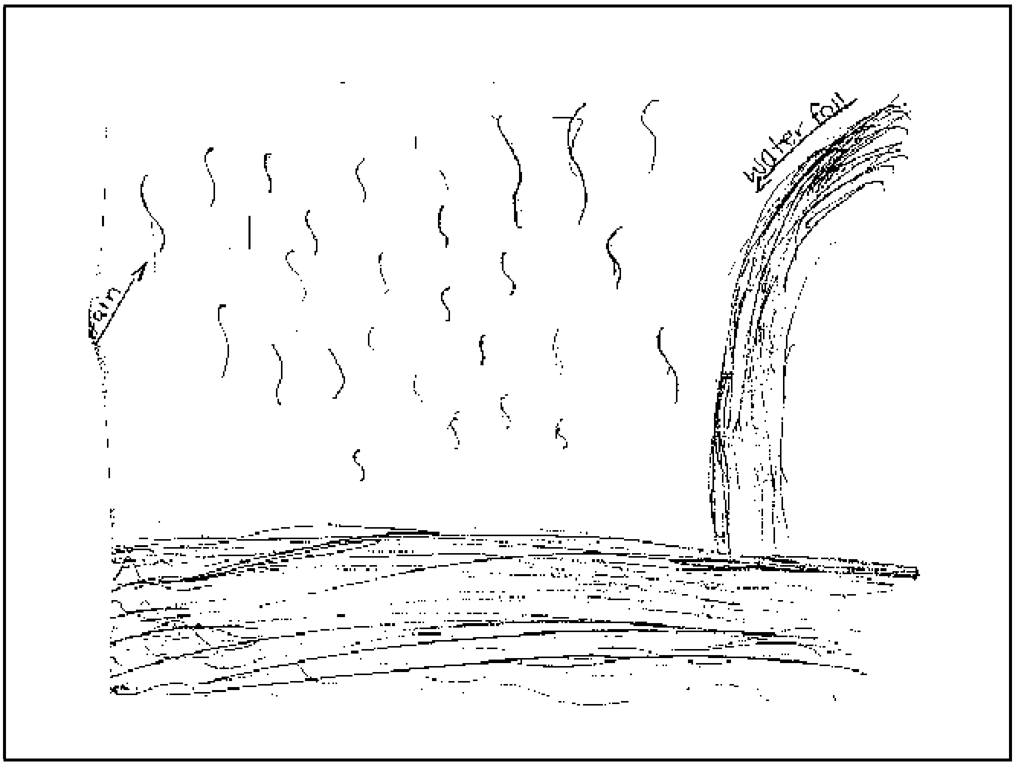
To be fully correct on the item shown on this page, the diagram had to include reference to all three of evaporation, transportation (e.g. of clouds by wind) and precipitation. Both of the responses shown on this page were assessed as fully correct, although one is a less conventional response than the other (and, strictly speaking, not factually correct—the important point is that the response refers to all three of the process aspects required).

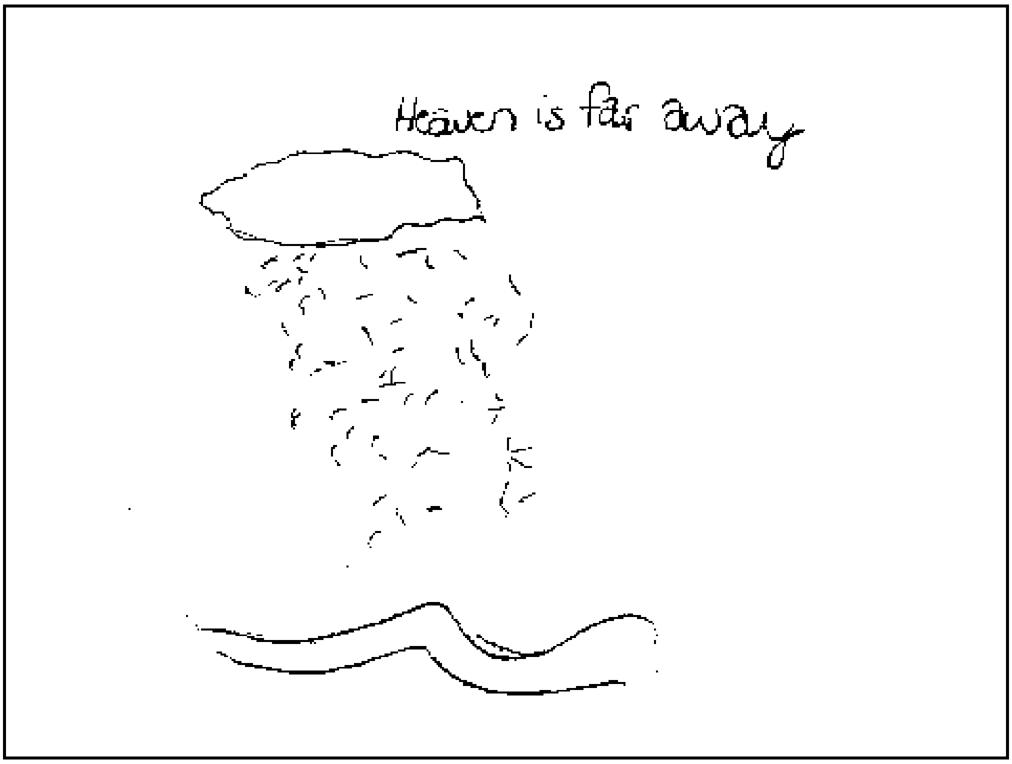
Draw a diagram to show how the water that falls as rain in one place may come from another place that is far away.



Aust: 59 %
Int'l: 57 %

The two responses shown on this page are examples of responses that could not be assessed as correct. The upper diagram received a partially correct mark for showing both a water source and precipitation, although the student obviously did not have much idea of the total process. The student who gave the response in the lower diagram had no idea of what is involved!





Some relationships with contextual factors

As mentioned earlier, data on a wealth of contextual variables were collected in TIMSS, by means of questionnaires to principals, teachers and students. Today I will report the relationship with achievement of some of the student-level variables only, for a mixture of variables that are often reported in other research studies and also for a few that are interesting in this particular context. To support the contention that the 'hands-on' performance assessment tasks in TIMSS measured rather different skills than were measured in the written tests, I will also provide for you some correlations between the two sets of tasks.

Student background variables

The relationship of a range of student demographic and other variables to mathematics and science achievement overall is shown for both populations in Table 2.

Table 2 Correlations between Student Background Variables and Achievement on TIMSS Written Tests

	Population 1		Population 2	
	Maths	Science	Maths	Science
Number of books in home	.18	.21	.27	.28
Family size	-.14	-.17	-.10	-.11
Parents' education status			.28	.30
Parents' occupation status	.27	.27	.30	.29
Home background composite	.36	.37	.40	.39
Language background composite	.11	.15	.10	.15
Word Knowledge	.61	.60	.47	.48
Like mathematics	.16	.15	.22	.10
Like science	.14	.16	.19	.25
Attribute success to luck	-.32	-.30	-.29	-.28
Self-efficacy belief	.16	.15	.32	.28

Because of the large sample sizes, these correlations are all highly significantly different from zero. With the exception of the associations between scores on the Word Knowledge test and achievement, the relationships indicated by the correlations in the table are not strong, though all are in the direction expected. The relatively high correlation between word knowledge and achievement was expected, particularly at Population 1, because the objective of embedding test items in contexts familiar to students meant that the mathematics and science items could not be divorced from reading, even though the language demands were kept as low as possible.

Number of books in the home, as an indicator of students' educational environment, has been found to be a useful predictor of achievement in earlier IEA studies. The 'home background composite' variable is an index built from the students' responses to the first four variables in the table, and was correlated only slightly more highly with achievement than three of the four variables on their own. Parents' countries of birth (whether English speaking or not English speaking, for one or both parents), and the extent to which English is spoken at home, were combined to form the 'language background composite' variable. A positive correlation indicates a greater degree of 'Englishness', which was found to be related to achievement but at a marginal level.

Among the affective variables measured in the Student Questionnaires, believing that success is due to luck or chance rather than to one's own efforts was the most predictive of achievement in all four tests, with higher levels of achievement associated with lower levels of such belief. 'Self-efficacy', or believing that one is doing well in the subject, was more highly related to achievement at Population 2 than at Population 1. Whether the students liked mathematics or science was only marginally associated with achievement at Population 1, slightly less marginally related at Population 2.

Correlations between some of the above student variables, with the addition of gender which was not related to achievement on the written tests in Australia, are shown in Table 3 for the 'hands-on' performance assessment tasks at Population 2. This part of the assessment was done by a much smaller number of students, although correlations of about .10 were still significantly different from zero at the .01 level of confidence.

Table 3 Correlations between Student Background Variables and Achievement on TIMSS 'Hands-on' Performance Assessment Tasks, Population 2

	Sex (1=F; 2=M)	Language (1=Eng, 2=not)	Word knowledge	Home bkgrd composite
Mathematics tasks				
Dice	.23	.12	.33	.33
Calculator	.11	-*	.42	.27
Folding & cutting	-	-.13	.23	.12
Around the bend	-	-	.41	.30
Packaging	.11	-.17	.40	.25
Plasticine	-	.21	.17	.19
Science tasks				
Pulse	.18	.16	.44	.29
Magnets	-	-	.44	.44
Batteries	-.18	-	.15	.22
Rubber band	.32	-	.29	.44
Solutions	.30	-	.36	.39
Shadows	-	-	.34	.27

* Correlation not significant at $p \leq .01$

Home background factors and word knowledge were mostly less important to achievement on the performance assessment tasks than on the written tests, especially tasks like 'Folding and cutting', 'Plasticine' and 'Batteries', which required a minimum of explanation of procedures followed or reasoning about conclusions. An English or non-English speaking background was not or only marginally associated with success on the performance assessment tasks—in favour of English speakers on some tasks and non-English speakers on other tasks. Apart from 'Rubber bands' and 'Solutions ', both of which required carrying out, recording results and justifying conclusions from an investigation, gender was only weakly associated with performance on any of the tasks. The only task performed better by boys than girls was the task requiring identification of which two of four batteries were worn out.

The final table of correlations, Table 4, shows the degree of association of success on each performance assessment task with achievement on the written tests, also at Population 2. The relationships are positive, as would be expected on measures of achievement, but generally lower than one would find between mathematics and science achievement assessed in written tests. The correlations are low enough to indicate that the performance assessment tasks are likely to be measuring somewhat different skills from those measured in the written tests. The correlations provide no clear justification for the proposed classification of some of the performance assessment tasks as 'mathematics' and some as 'science ', however.

Table 4 Correlations of Performance Assessment Tasks with Written Tests, Population 2

	Written test total score	
	Mathematics	Science
Mathematics tasks		
Dice	.35	.34
Calculator	.36	.36
Folding & cutting	.28	.27
Around the bend	.43	.46
Packaging	.45	.41
Plasticine	.25	.25
Science tasks		
Pulse	.44	.47
Magnets	.44	.42
Batteries	.26	.27
Rubber band	.41	.31
Solutions	.45	.38
Shadows	.43	.38

I would also like to share with you some graphs of relationships to achievement of some variables that are often thought to be 'culprits' in partly accounting for lower achievement and some that might be expected to be associated with higher achievement. The variables are 'extent of daily TV watching' (Figure 1), 'time spent each day in playing computer games' (Figure 2), 'time spent each day on homework' (Figure 3), and 'frequency of students' doing science experiments' (Figure 4). Relationships were usually similar at both Population levels.

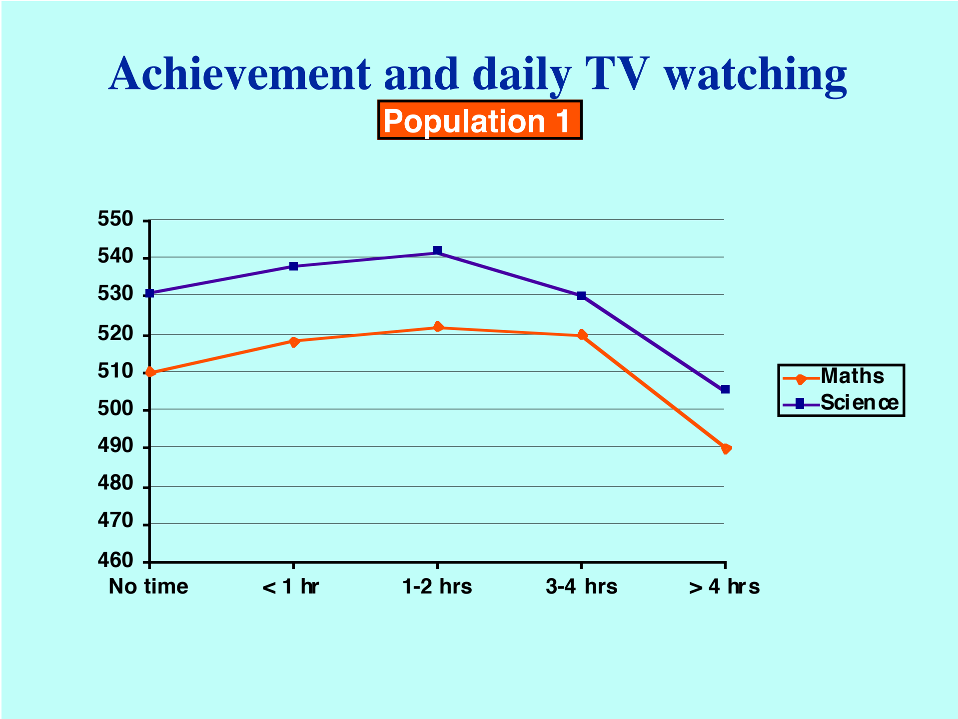


Figure 1 Mathematics and Science Achievement and Time Spent Watching TV

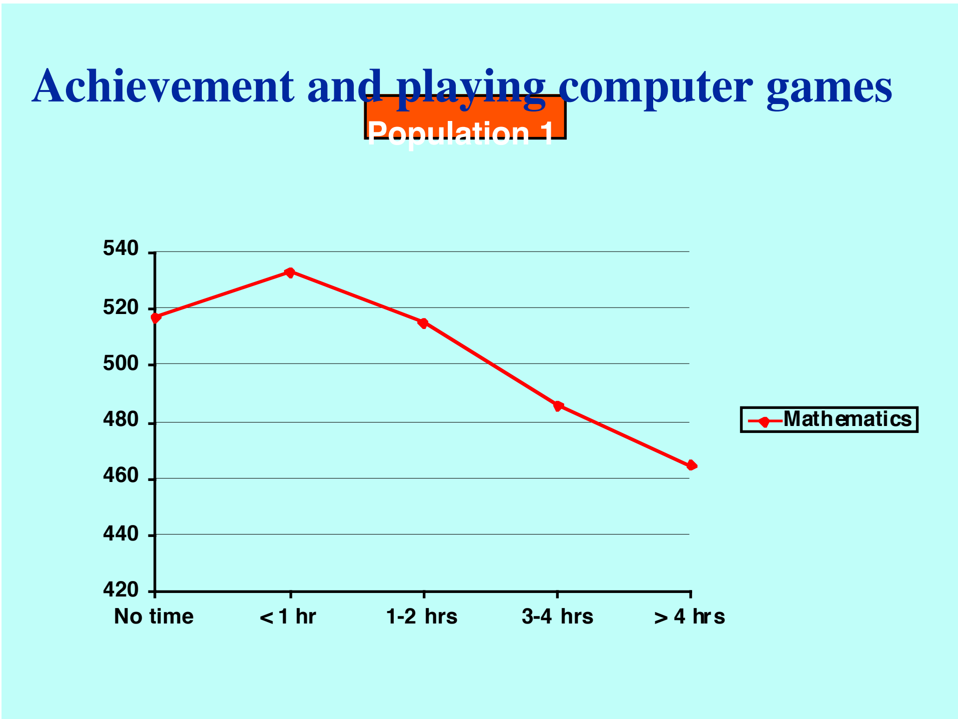


Figure 2 Mathematics Achievement and Time Spent Playing Computer Games

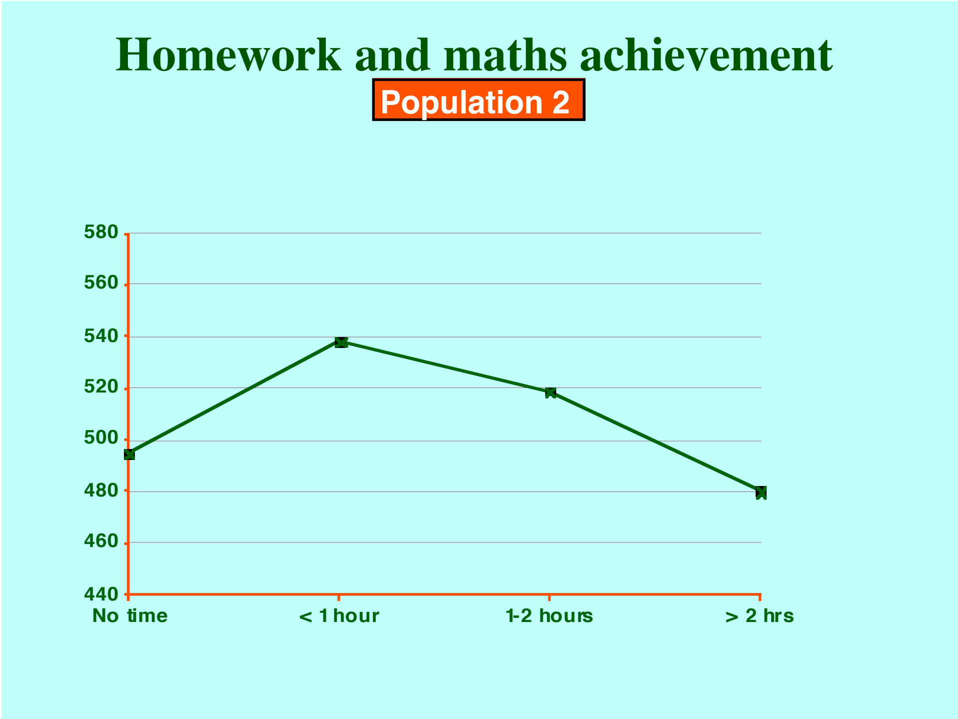


Figure 3 Mathematics Achievement and Daily Time Spent on Homework

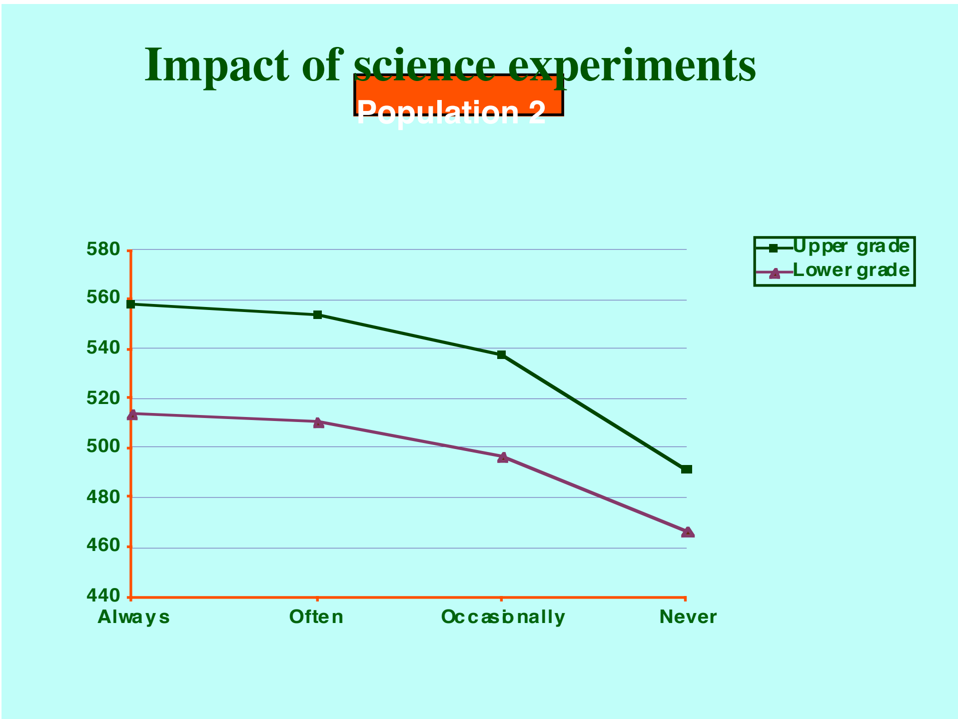


Figure 4 Science Achievement and How Often Experiments are Done in Science Lessons

As we might expect, there are some detrimental effects, as far as school achievement is concerned, associated with too much watching of TV or playing of computer games. For both, though, there is an initial rise in achievement for students who spend only a moderate amount of time—up to two hours a day (TV) or one hour a day (games). After that, achievement declines, though not seriously until beyond four hours a day of TV and two hours a day of computer games. We cannot attribute cause and effect here—the association with lower achievement is clear, but it may be that the less able students are drawn to more TV watching and game playing, rather than that the large amount of time spent in these activities is the reason for their lower achievement. A similar comment can be made about the relationship between amount of homework and achievement, an example of which is shown in Figure 3. One suspects that the lower achievement of those who say they do the most homework has more to do with the students' ability level than with the time spent.

Finally, a classroom practice variable that is positively associated with achievement on the TIMSS science test is illustrated in relation to achievement in Figure 4. The relationship is positive, because the more often students do experiments in their science classes (as reported by the students themselves), the better their achievement. The largest differences in achievement occur for the steps from 'never' to 'occasionally', and from 'occasionally' to 'often'.

IN CONCLUSION

The results I have shared with you today are only a tiny portion of the information and insights that we can glean from the large amount of data collected in TIMSS. Others presenting papers at this conference today and tomorrow will contribute further insights from the Australian data, from international data, and from other related initiatives that have taken place in Australia recently in the fields of numeracy and science education.

We have seen that the Australian students' performance was excellent in some areas, poor in others—especially in some aspects of mathematics. We have also seen that our students performed very creditably on the world stage, though not quite at the very top level. I should like to close with some questions that we need to reflect on, as policy makers, researchers, and mathematics and science educators, as we determine the kind of mathematics and science education Australian students will receive in the years to come. Most of these could arise without reference to TIMSS, but the experience and findings from TIMSS reinforce their importance:

- Is sufficient curriculum time being devoted to mathematics and science instruction in Australia, particularly at primary level? Are there priorities *within* mathematics and science that should be changed? If so, what are these, and how might we go about making the changes?

- Are our mathematics and science curricula too broad? Or perhaps too narrow? Are we introducing topics later than we need to, in terms of what students are capable of doing? Are we doing enough to cater for the interests and capacities of our best students?

- How can we best make use of technological aids (e.g. calculators and computers) in our mathematics and science teaching? Are there dangers in introducing such aids too early, so that students become unduly reliant on them? Have our teachers had sufficient opportunities to adapt their teaching to incorporate the best uses of computers and calculators? What *are* the best uses?

- Do we need to rethink the nature and role of homework, practice and self-study? How important is feedback to students on the homework they have done?

- Are we equipping students with appropriate skills to cope with and make critical judgements about the large amounts of information they now have access to? Are they 'numerate' enough, for example, to judge advertising material?

- How important are the skills of being able to explain, justify or generalise? TIMSS suggests that students in all countries have trouble with these. Should we be facilitating the development of these skills more than we are doing now?

- Finally, where would we want to be placed in the next major cross-national study? What would we need to do to ensure that we can achieve that goal?

We hope that the discussion sessions we have set up for this conference will be useful in addressing these questions, and others you may wish to raise.

---

[i] J. Lokan, P. Ford & L. Greenwood, *Maths and Science On the Line: Australian Junior Secondary Students' Performance in the Third International Mathematics and Science Study*, TIMSS Australia Monograph No. 1, Melbourne, ACER, 1996; and

J. Lokan, P. Ford & L. Greenwood, *Maths and Science On the Line: Australian Middle Primary Students' Performance in the Third International Mathematics and Science Study*, TIMSS Australia Monograph No. 2, Melbourne, ACER, 1997.