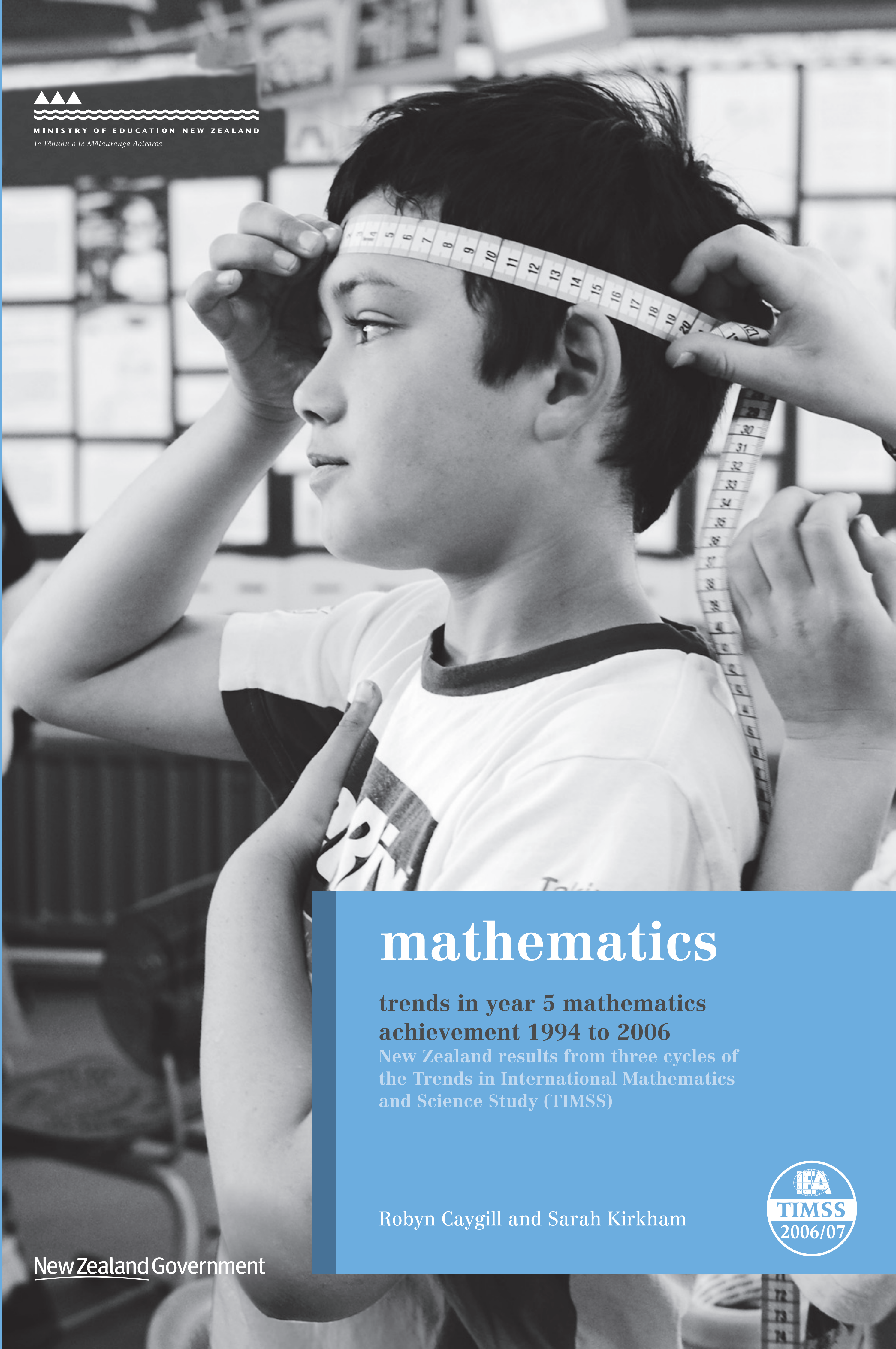
TIMSS December 2008

MINISTRY OF EDUCATION NEW ZEALAND
Te Tāhuhu o te Mātauranga Aotearoa



mathematics

trends in year 5 mathematics achievement 1994 to 2006

New Zealand results from three cycles of the Trends in International Mathematics and Science Study (TIMSS)

Robyn Caygill and Sarah Kirkham

IEA TIMSS 2006/07

New Zealand Government

Overview of TIMSS

What is TIMSS?

The Trends in International Mathematics and Science Study (TIMSS) measures trends in mathematics and science achievement at the fourth and eighth grades (Years 5 and 9) as well as monitoring curricular implementation and identifying the most promising instructional practices from around the world.

Conducted on a regular 4-year cycle, TIMSS has assessed mathematics and science in 1994/95[1], 1998/99, 2002/03, and 2006/07 with planning underway for 2010/11.

What does TIMSS consist of?

TIMSS consists of assessments of students' achievements in mathematics and science along with questionnaires for students, teachers, and principals to gather background information. The background information provides a context within which the achievement can be examined.

The TIMSS assessments are organised around two dimensions: a content dimension specifying the domains or subject matter to be assessed within mathematics and science; and a cognitive dimension specifying the domains or thinking processes to be assessed. These domains are published in the *TIMSS 2007 assessment frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, and Erberber, 2005). To guide questionnaire development, the contextual factors associated with students' learning in mathematics and science are also included in the frameworks.

How was TIMSS developed?

The TIMSS tests were developed cooperatively with representatives from those participating countries that have been involved throughout the entire process. Questions were field-tested with a representative sample of students in these countries and the results generated were used to select and refine the questions for the final test. Questions for the background questionnaires underwent a similar process.

Who participated?

In TIMSS 2006/07, approximately 425,000 students in 59 countries from all around the world took part. Participants included 183,150 students from 37 countries and 7 benchmarking participants at the middle primary level, and 241,613 students from 50 countries and 7 benchmarking participants at the lower secondary level.[2] In this cycle of TIMSS, only Year 5 students from New Zealand participated.

Who administered TIMSS?

A consortium was responsible for managing the international activities required for the project. This consortium comprised: the International Study Centre, Lynch School of Education at Boston College, (Massachusetts) United States; the IEA Secretariat in Amsterdam, the Netherlands; the IEA's Data Processing Centre in Hamburg, Germany; Statistics Canada in Ottawa, Canada; and the Educational Testing Service (ETS) in Princeton, New Jersey in the United States. In New Zealand the Comparative Education Research Unit in the Ministry of Education was responsible for carrying out TIMSS.

What procedures were used to ensure the quality of the data?

TIMSS procedures are designed to ensure the reliability, validity, and comparability of the data through careful planning and documentation, cooperation among participating countries, standardised procedures, and attention to quality control throughout. Procedures included verification of translations and layout of booklets and questionnaires, monitoring of sampling activities, international and national quality control observers during test administration, checking of data, detailed manuals covering procedures, and rigorous training for all involved. Members of the consortium ensured procedures were adhered to by all participating countries.

Why participate in TIMSS?

Although it is often assumed that the international studies are only useful for international benchmarking purposes, the real value of TIMSS lies in its ability to provide a rich picture of mathematics and science achievement within New Zealand and over time.

TIMSS (along with other international assessment studies) can provide information about the performance of the New Zealand education system at the national level within a global context. The information from studies such as TIMSS is used in the development and review of policy frameworks and also to inform and improve teaching practice. Developments arising out of previous cycles of TIMSS include resource materials for schools and teachers along with teacher in-service training programmes.

1 Note that this cycle of the study is called TIMSS 1995 internationally as most countries participated in 1995. However southern hemisphere countries conducted the assessment towards the end of 1994 so in New Zealand reports the study is referred to as TIMSS 1994/95. Similarly for the subsequent cycles, the two years in which administrations occurred in participating countries are indicated.

2 Mongolia does not appear in any international comparisons because they were unable to meet sampling criteria. Selected results for Mongolia appear in Appendix E of Martin, Mullis, and Foy (2008). Throughout the report 36, rather than 37, countries are discussed at the middle primary level.

Table of Contents

Acknowledgements

TIMSS is a collaborative effort internationally and nationally. There are many people that have been involved and it is not possible to thank all of them individually here. However, there are a few key people that we would like to acknowledge.

This study was made possible by the cooperation of the schools, teachers, and students that participated. Thanks to these participants, large amounts of data have been collected about the teaching and learning of New Zealand Year 5 students.

Thanks also to the many members of the Research Division, and also those from the wider ministry, that have contributed to the successful collection of data for TIMSS, particularly Kate Lang. Special thanks are also extended to Megan Chamberlain and Ian Schagen for their help with the technical aspects of analysing the data.

Finally thanks to those that contributed to this publication by reviewing, editing and publishing this report.

Robyn Caygill
TIMSS National Research Coordinator

Sarah Kirkham

Key findings

Achievement

- There has been a significant overall improvement since the first TIMSS cycle in 1994/95. However, New Zealand's mean mathematics achievement in 2006/07 was not significantly[3] different from TIMSS 2002/03.
- The proportions of New Zealand students reaching the advanced, high, intermediate, and low benchmarks have not changed significantly since TIMSS 2002/03.
- New Zealand Year 5 students, on average, achieved above the mean mathematics achievement for 12 of the 36 countries that participated in TIMSS 2006/07 at the middle primary level.
- Relative to other countries, as well as Year 5 students' overall performance, New Zealand students were stronger on *data display* questions and relatively weak on *number* questions. Students also performed relatively better on questions that involved *reasoning* compared to *knowing* or *applying*.

Background context

- There was no difference between mean mathematics achievement of boys and girls in TIMSS 2006/07. Both boys and girls have shown a significant improvement since 1994.
- Both high and low performers were found in all ethnic groupings. Asian and Pākehā/European students demonstrated significantly higher mean mathematics scores than Māori and Pasifika students. Asian students performed significantly higher than Pākehā/European students. Māori students performed significantly higher than Pasifika students.
- Students who always or almost always spoke English at home had higher mathematics achievement, on average, than those who sometimes or never spoke English at home. Students who were born in New Zealand had higher mathematics achievement, on average, than those who were not.
- Students from higher socio-economic backgrounds tended to have higher mean mathematics achievement than those from lower socio-economic backgrounds as evidenced by the proxy measures *books in the home*, *items in the home*, *household size* and *mobility*. In addition, the decile of the school they attended, indicative of the level of economic disadvantage in the community in which they live, was positively related to mathematics achievement.

Student attitudes

- New Zealand middle primary students were generally positive towards mathematics. The proportion has not changed since TIMSS 1994/95. Students who were more positive towards mathematics had, on average, higher achievement than those who were more negative.
- About half of the New Zealand Year 5 students in TIMSS 2006/07 expressed a high level of self-confidence in mathematics.
- About the same proportions of Year 5 girls and boys were very positive about mathematics. However, proportionally more boys than girls in New Zealand expressed a high level of self-confidence in mathematics.
- Proportionally more Asian students reported positive attitudes to mathematics and fewer Pākehā/European reported positive attitudes to mathematics. Māori and Pasifika students expressed lower self-confidence in mathematics compared with students in the Pākehā/European, Asian, and Other ethnic groupings.

3 The term 'significantly' is used throughout this report to refer to statistical significance.

Introduction

This report examines the mathematics results for New Zealand Year 5 students from TIMSS 2006/07.[4] Along with the report on New Zealand's results for science (Caygill 2008), this report forms the beginning of a series of reports around New Zealand's participation in TIMSS 2006/07. International findings for mathematics for TIMSS 2006/07 have been published by the IEA (Mullis, Martin & Foy, 2008). A separate international report on science was also published at this time (Martin, Mullis & Foy, 2008).

This report begins by examining trends in New Zealand mathematics achievement at the Year 5 level from 1994 to 2006. It then looks at New Zealand's mathematics achievement in relation to other countries that participated in the study. An examination of the TIMSS assessment questions in relation to New Zealand's mathematics curriculum is presented along with analyses of achievement by sub-groupings (such as gender and ethnicity) and background factors. Lastly, a statistical model that attempts to explain variations among students, classes, and schools, using the background information discussed in this report is also described.

Assessment of mathematics in TIMSS

The TIMSS assessment has two main dimensions: a dimension that describes the content or subject matter to be assessed; and a dimension that describes the cognitive processes used to answer the questions. The three content dimensions for mathematics are: number, geometric shapes and measures, and data display. The detail about the topic areas covered in these domains at each grade or year level assessed and a set of assessment objectives for each topic area are presented in the *TIMSS 2007 assessment frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, and Erberber, 2005). Briefly, each of the content areas is described in the frameworks as follows.

"The number content domain for the fourth grade includes understanding of place value, ways of representing numbers, and the relationships between numbers." (p. 16).

"The geometric shapes and measures domain includes properties of geometrical figures such as lengths of sides, sizes of angles, areas, and volumes." (p. 19).

"The data display content domain includes reading and interpreting displays of data." (p. 21).

In order to answer questions in the TIMSS test correctly, as well as being familiar with the mathematics content, students need to draw on a range of cognitive skills. Also, in their lives outside and beyond school, students will need to do more than accurately recall a range of mathematic facts. This is acknowledged in the framework with three aspects to the cognitive dimension entitled knowing, applying, and reasoning. Briefly, each cognitive dimension is described in the framework as follows.

4 Internationally this cycle of the study is called TIMSS 2007. As southern hemisphere countries conducted the study first, towards the end of 2006, it is referred to as TIMSS 2006/07 throughout this report.

"The first domain, knowing*, covers the facts, procedures, and concepts students need to know, while the second,* applying*, focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions. The third domain,* reasoning*, goes beyond the solution or routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems."* (p. 33).

Data collection

Each student was assessed in two timed sessions of 36 minutes, and answered a combination of mathematics and science questions. The assessment was a pencil-and-paper test containing both multiple-choice and constructed-response questions. Following this, students were given a questionnaire containing questions about themselves, their opinions about mathematics and science, their computer use and time spent on homework. Principals and teachers were also given questionnaires in order to gain further information about the context in which the mathematics teaching and learning take place. In New Zealand, the assessments and questionnaires were conducted in English.[5]

International participants in TIMSS

The number of participants in TIMSS at the Year 5 or grade 4 level has steadily increased since 1994, when 26 countries took part. In 2002, 25 countries and 3 benchmarking participants took part. Benchmarking participants are usually states or parts of countries and are not included in international averages. In 2006, the number of education systems participating at the middle primary level had risen to 37 countries and 7 benchmarking participants.

Technical information

A lot of information is gathered during the TIMSS administration and a number of techniques are applied when collecting and analysing the data. The *TIMSS 2007 technical report* (Olson, Martin, & Mullis (Eds.), 2008) contains a detailed account of the procedures for scoring, translation of materials, sampling, survey operations, quality assurance, sampling weights, item analysis, scaling, and reporting. In addition, the *TIMSS 2007 user guide for the international database* (to be published in early 2009) contains information on how to analyse the data. Brief details of the technical information are given in the definitions and technical notes at the end of this report.

TIMSS encyclopaedia

In order to provide a context in which the TIMSS results can be examined, TIMSS also publishes the *TIMSS 2007 encyclopedia: a guide to mathematics and science education around the world* (Mullis, Martin, Olson, Berger, Milne, & Stanco (Eds.), 2008). This encyclopaedia contains short reports from each country describing mathematics and science education policies and practices in that country.

5 In 2002, tests and questionnaires were also translated into te reo Māori, but in order to make comparisons between each of the cycles, these students were excluded from analyses presented in this report.

Trends in New Zealand mathematics achievement 1994 to 2006

Trends in means and ranges since 1994

New Zealand has participated in TIMSS since its inception in 1994. In 1998, although no assessment was offered internationally at the middle primary level, New Zealand opted to repeat the 1994 assessment. Therefore, we now have information from four different assessments of mathematics achievement. Figure 1 presents the distributions of mathematics achievement of New Zealand Year 5 students over the four cycles of TIMSS.

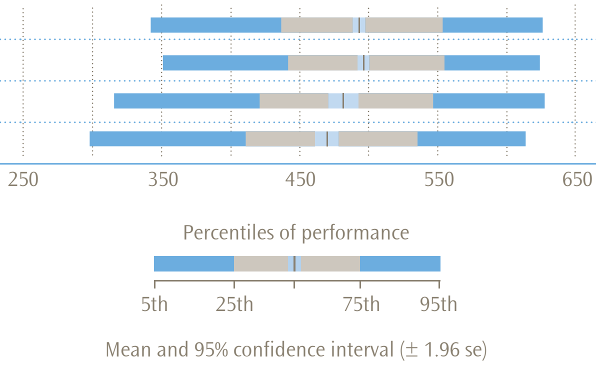
The results from an examination of mathematics achievement since 1994 (see Figure 1) show that mean mathematics achievement in 2006 is higher than 1994, the first cycle of TIMSS. Although the mean score for 2006 is numerically lower than 2002, the difference between 2002 and 2006 is not significant.

It is also useful to look at the range of achievement as represented by the outer limits of achievement. The lowest outer limit presented in Figure 1 is the 5th percentile – the score at which only five percent of students achieved a lower score and 95 percent of students achieved a higher score. The highest outer limit is the 95th percentile – the score at which only five percent of students achieved a higher score and 95 percent of students a lower score. In addition, the 25th and 75th percentiles are presented in Figure 1, along with the inter-quartile range.

As shown in Figure 1, the range of achievement was narrower in 2006 than 1998 and 1994, but not as narrow as in 2002. The positive aspect of this change is that fewer students are demonstrating very low achievement, while a similar proportion of New Zealand students are gaining very high scores.

Figure 1 **Distribution of New Zealand Year 5 mathematics achievement in TIMSS from 1994 to 2006**

Year	Mean mathematics score		Distribution of mathematics achievement	Range from 5th to 95th percentile	Inter-quartile range from 25th to 75th percentile
2006	492	(2.3)		284	117
2002	496	(2.1)		273	114
1998	481	(5.6)		312	126
1994	469	(4.4)		316	125



Note: For trend purposes, only students tested in English are included in the results for 2002. Standard errors are presented in parentheses.

Trends in benchmarks for mathematics

In order to describe more fully what achievement on the mathematics scale means, the TIMSS international researchers have developed benchmarks. These benchmarks link student performance on the TIMSS mathematics scale to performance on the mathematics questions and describe what students can typically do at set points on the mathematics achievement scale. The international mathematics benchmarks are four points on the mathematics scale; the advanced benchmark (625), the high benchmark (550), the intermediate benchmark (475), and the low benchmark (400). The performance of students reaching each benchmark is described in relation to the types of questions they answered correctly. Table 1 presents the descriptions of the international benchmarks of mathematics achievement.

Table 1 **TIMSS 2006/07 international benchmarks of mathematics achievement**

Advanced international benchmark – 625

Students can apply their understanding and knowledge in a variety of relatively complex situations and explain their reasoning. They can apply proportional reasoning in a variety of contexts. They demonstrate a developing understanding of fractions and decimals. They can select appropriate information to solve multi-step word problems. They can formulate or select a rule for a relationship. Students can apply geometric knowledge of a range of two- and three-dimensional shapes in a variety of situations. They can organize, interpret, and represent data to solve problems.

High international benchmark – 550

Students can apply their knowledge and understanding to solve problems. Students can solve multi-step word problems involving operations with whole numbers. They can use division in a variety of problem situations. They demonstrate understanding of place value and simple fractions. Students can extend patterns to find a later specified term and identify the relationship between ordered pairs. Students show some basic geometric knowledge. They can interpret and use data in tables and graphs to solve problems.

Intermediate international benchmark – 475

Students can apply basic mathematical knowledge in straightforward situations. Students at this level demonstrate an understanding of whole numbers. They can extend simple numeric and geometric patterns. They are familiar with a range of two-dimensional shapes. They can read and interpret different representations of the same data.

Low international benchmark – 400

Students have some basic mathematical knowledge. Students demonstrate an understanding of adding and subtracting with whole numbers. They demonstrate familiarity with triangles and informal coordinate systems. They can read information from simple bar graphs and tables.

Source: Exhibit 2.1 from Mullis, Martin, and Foy, 2008.

Table 2 presents the proportions of New Zealand Year 5 students that reached each of the benchmarks in each cycle from 1994 to 2006. Note that the proportion shown for the low benchmark also includes students who performed at the advanced, high, and intermediate benchmarks. This is because, by definition, students who could do the more complex questions associated with, for example, the high benchmark, would also be able to complete the easier questions associated with the intermediate and low benchmarks.

The proportion of New Zealand students reaching the high, intermediate, and low benchmarks has been steadily rising since 1994 although the advanced benchmark has not changed significantly during this time. Five percent of students reached the advanced benchmark in 2006. The proportion of students reaching the high, intermediate and low benchmarks, which peaked in 2002 (27%, 62%, and 86% respectively), has been maintained in the 2006 results. The differences between the benchmarks in 2006 compared with 2002 are not of statistic significance.

There was also a group of Year 5 students in each cycle who did not reach the low benchmark. In terms of the benchmark definitions, these were students who did not demonstrate some basic mathematical knowledge. This group was proportionally largest in 1994 (22%) and smallest in 2002 (14%).

Table 2 **Trends in proportions of Year 5 students at each benchmark from 1994 to 2006**

	Percentage of Year 5 students reaching each benchmark			
Year	Advanced	High	Intermediate	Low
2006	5 (0.5)	27 (1.0)	61 (1.1)	85 (1.0)
2002	5 (0.5)	27 (1.2)	62 (1.3)	86 (1.0)
1998	5 (0.9)	24 (1.9)	55 (2.5)	81 (1.8)
1994	4 (0.6)	20 (1.4)	51 (1.9)	78 (1.7)

Note: Standard errors are presented in parentheses.

Trends on the test questions

At the end of each cycle of TIMSS, test questions are released into the public domain. At the beginning of the next cycle, new questions are developed to replace the released questions. In addition, each cycle of TIMSS includes some questions from the previous cycle(s) to provide a trend measure over time. This section presents an analysis of the trend questions included in both TIMSS 2002/03 and TIMSS 2006/07. Note that no questions from TIMSS 1994/95 were included in the TIMSS 2006/07 assessment.

There were 79 questions common to both the 2002/03 and 2006/07 cycles. Of these 79 questions, 10 questions had similar proportions of students correctly answering them across two cycles (as shown in Table 3). There were a number of questions (39) that proportionally fewer students correctly answered in 2006 compared with 2002. In contrast, there were 30 questions that proportionally more students correctly answered in 2006 compared with 2002. When the change in proportions of students correctly answering was averaged across all the common questions, this represented a decrease of 0.6 percent.

This analysis reiterates that the decrease of New Zealand's mean mathematics score by 3 scale score points (from 495 to 492) is not statistically significant.

Table 3 **Trends in question statistics for mathematics questions common to 2002/03 and 2006/07**

Change between 2002/03 and 2006/07	Decrease by 5% or more	Decrease by between 1% and 5%	Increase or decrease by 1% or less	Increase by between 1% and 5%	Increase by 5% or more
Number of questions	14	25	10	21	9

It is interesting to note that of the 14 questions in the group that decreased by 5 percent or more (when the proportion of students correctly answering in 2006 was compared with 2002), there were proportionally more of the geometric shapes and measures, and data display questions than number questions. In contrast, proportionally more questions from the number domain were in the group that increased and far fewer geometric shapes and measures questions.

Trends in mathematics content and cognitive domains

The mathematics assessment in TIMSS is organised around two dimensions, a content dimension and a cognitive dimension, as described in the *TIMSS 2007 assessment frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, and Erberber, 2005). The content dimension comprises three content domains that describe the subject matter to be assessed:

- number;
- geometric shapes and measures; and
- data display.

The three content domains can be mapped onto the three strands of the current New Zealand Mathematics and Statistics curriculum, Number and Algebra, Geometry and Measurement, and Statistics, which themselves are combinations of the five strands of the previous Mathematics curriculum (Number, Algebra, Measurement, Geometry, and Statistics).

The cognitive dimension comprises three cognitive domains that describe the thinking processes that students must use as they engage with the content:

- knowing;
- applying; and
- reasoning.

TIMSS assessment questions were categorised by the content and cognitive domains, and content and cognitive achievement scales were constructed separately for each domain. In order to simplify comparisons across domains, the scales were constructed to have the same average difficulty (set at 500 scale score points). As well as looking at achievement in each of these domains, the results can then be used to ascertain relative strengths for participating countries.

As Table 4 shows, New Zealand Year 5 students achieved relatively better at data display questions and relatively worse at number questions in 2006. Although the 2002 domains of *number* and *patterns and relationships* have been combined in 2006, and likewise the 2002 domains of *measurement* and *geometry* have been combined, the trend is very similar. In 2002, the *data* domain (522) was relatively higher than the other domains and the *number* (475) and *patterns and relationships* (495) domains were relatively lower than the other domains.

In the cognitive domains, New Zealand Year 5 students achieved relatively better at tasks that required them to use their reasoning and relatively worse at questions that required demonstrating their knowledge in 2006. In 2002, New Zealand year 5 students showed a relative strength in the reasoning domain (503) and a relative weakness in the applying domain (486).

Table 4 Year 5 mean mathematics scores on the content and cognitive domains in 2006

Content domain	Mean domain score	Cognitive domain	Mean domain score
Number	478 (2.7)	Knowing	482 (2.5)
Geometric shapes and measures	502 (2.3)	Applying	495 (2.3)
Data display	513 (2.6)	Reasoning	503 (2.8)

Note: Standard errors are presented in parentheses.

Table 5 shows the number of test questions (and the associated raw score points) in each of the content and cognitive domains. As can be seen from the table, score points were not evenly distributed across domains. This distribution of questions across domains reflects the content and cognitive emphasis of many of the curricula of participating countries.

Looking at Tables 4 and 5 together, it is important to note that the content domain where New Zealand Year 5 students show the greatest strength, data display, had the least number of questions. Similarly, the cognitive area of greatest strength, reasoning, had the least number of questions. The distribution of mathematics questions across the content domains was similar in 2006 to 2002, with a slight increase in data display questions (*data* in 2002) and corresponding decrease in number questions (*number* and *patterns and relationships* in 2002).

Table 5 Number of questions in each of the content and cognitive domains

Content domain	Total number of questions	Total number of score points
Number	93	98
Geometric shapes & measures	60	65
Data display	26	29
Total	**179**	**192**

Cognitive domain	Total number of questions	Total number of score points
Knowing	69	73
Applying	70	75
Reasoning	40	44
Total	**179**	**192**

Note: In scoring the tests, correct answers to most questions were awarded one point. However, responses to some constructed-response questions were evaluated for partial credit with a fully correct answer awarded two points. Thus, the number of score points exceeds the number of questions in the test.

New Zealand mathematics achievement in 2006 in an international context

As shown in Figure 2, the mean mathematics score for New Zealand Year 5 students in TIMSS 2006/07 was 492 scale score points. New Zealand's mean score was similar to the Czech Republic (486), Scotland (494), the Slovak Republic (496) and Armenia (500), and significantly higher than 12 other countries. In contrast, 19 countries had higher mean mathematics achievement, including Singapore (599), England (541), the United States (529) and Australia (516).

The range of achievement (from the 5th to 95th percentile) in New Zealand was 284 score points from 341 (the 5th percentile) to 626 (the 95th percentile)[6]. This was relatively wider than the ranges of many of the higher achieving countries but much the same as that of England (284). Another measure of spread, the inter-quartile range (from the 25th to 75th percentile) can also be examined. For New Zealand (117) this was relatively wider than the higher performing countries.

Given the number of countries now participating in TIMSS, it is more meaningful to compare New Zealand to a selection of countries (such as English-speaking or high-performing). Compared to the countries that tested in English (Singapore, England, the United States, Australia, and Scotland), New Zealand had significantly lower mathematics achievement, on average, than all of them except for Scotland (numerically New Zealand was lower but the difference is not significant).

Alongside Figure 2, Table 6 presents some information to help put mathematics achievement in context. Countries are presented in the same order as in Figure 2. It contains information on the number of years of primary schooling students will have undertaken by the time of the assessment, along with students' average age at the time of testing. Also given in the table is the average number of hours of time spent in mathematics instruction during the assessment year according to teacher reports. Three bits of information are presented about the economic circumstances, on average, across each country, the Human Development Index, and two versions of the Gross National Income per Capita (described later).

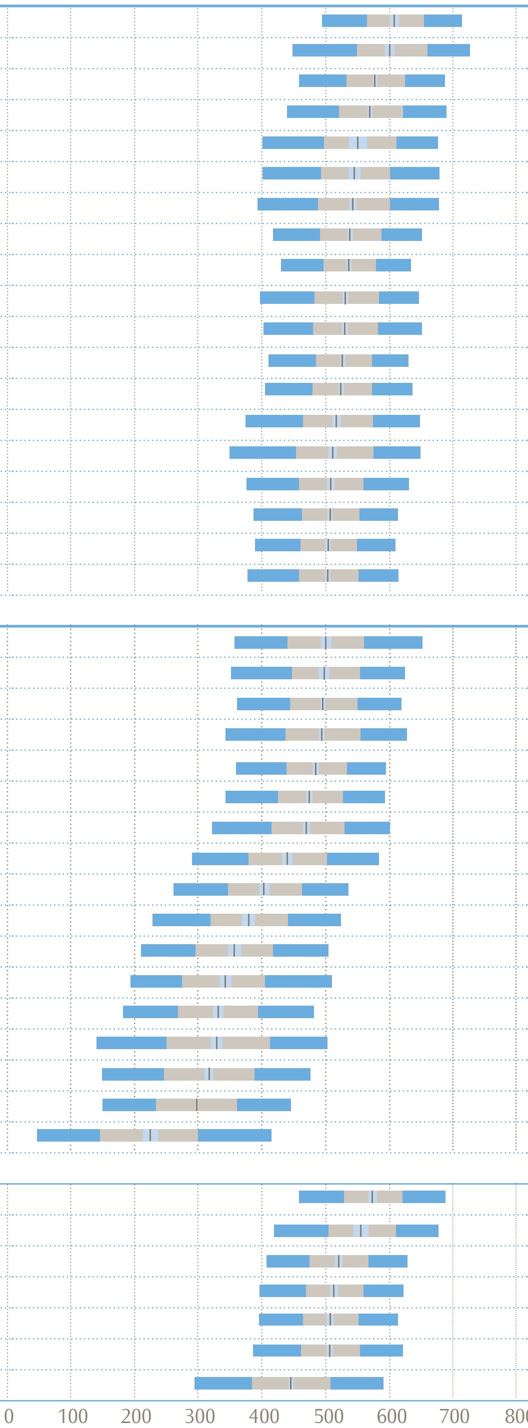
New Zealand spends less time at the middle primary level teaching mathematics, on average, according to teacher reports, than any of the other English-speaking countries. However, New Zealand did spend more time on mathematics instruction than the top-performing non-English speaking countries with the exception of Hong Kong SAR. Teachers in Hong Kong SAR reported spending a similar number of mathematics instructional hours as New Zealand.

6 Because results are rounded to the nearest whole number, this difference may appear inconsistent.

Table 6 also presents the Human Development Index provided by the United Nations Development Programme (UNDP – for details see *Human Development Report 2007/2008*, p. 229-232). This index was included by Mullis, Martin, and Foy (2008) in the international reporting to provide some context around the economic and educational development of TIMSS participating countries. The index ranges from a minimum value of 0 to a maximum value of 1, with high values indicating that people in a country generally enjoy long life expectancy, high levels of school enrolment and adult literacy, and a good standard of living as measured by per capita GDP. New Zealand was relatively high on this scale with a value of 0.943, similar to that of Italy (0.941), and England and Scotland (0.946 – this value is actually for the United Kingdom as no disaggregated data is available for England and Scotland) and lower than that of Australia (0.962) and the United States (0.951).

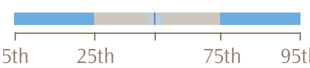
Perhaps easier to relate to than the HDI, two versions of the Gross National Income (GNI) per Capita are also presented in Table 6. The first of the two columns gives the GNI per Capita in United States dollars while the second is an adjusted value that takes account of comparative purchasing power between each country and the United States. Compared to the countries that assessed in English, New Zealand has the lowest income regardless of which of these values is used.

Figure 2 Distribution of middle primary mathematics achievement in TIMSS 2006/07



	Country	Average scale score		Range 5th to 95th percentile	Inter-quartile range 25th to 75th percentile
	Hong Kong SAR	607 (3.6)	▲	219	88
	Singapore	599 (3.7)	▲	278	111
	Chinese Taipei	576 (1.7)	▲	229	92
	Japan	568 (2.1)	▲	250	100
1	Kazakhstan	549 (7.1)	▲	275	114
	Russian Federation	544 (4.9)	▲	277	108
	England	541 (2.9)	▲	284	113
1	Latvia	537 (2.3)	▲	234	96
**	Netherlands	535 (2.1)	▲	203	82
1	Lithuania	530 (2.4)	▲	249	101
2 *	United States	529 (2.4)	▲	249	102
	Germany	525 (2.3)	▲	219	88
*	Denmark	523 (2.4)	▲	231	93
	Australia	516 (3.5)	▲	274	109
	Hungary	510 (3.5)	▲	300	122
	Italy	507 (3.1)	▲	255	101
	Austria	505 (2.0)	▲	226	90
	Sweden	503 (2.5)	▲	220	88
	Slovenia	502 (1.8)	▲	237	93
	TIMSS scale avg.	**500**			
	Armenia	500 (4.3)		295	120
	Slovak Republic	496 (4.5)		273	108
*	Scotland	494 (2.2)		259	107
	New Zealand	**492 (2.3)**		**284**	**117**
	Czech Republic	486 (2.8)		236	95
	Norway	473 (2.5)	▼	250	102
	Ukraine	469 (2.9)	▼	278	114
1	Georgia	438 (4.2)	▼	293	123
	Iran, Islamic Rep. of	402 (4.1)	▼	274	116
	Algeria	378 (5.2)	▼	295	121
	Colombia	355 (5.0)	▼	295	122
	Morocco	341 (4.7)	▼	316	130
	El Salvador	330 (4.1)	▼	300	126
	Tunisia	327 (4.5)	▼	362	162
	Kuwait	316 (3.6)	▼	327	142
	Qatar	296 (1.0)	▼	296	127
	Yemen	224 (6.0)	▼	368	153
Benchmarking participants					
2	Massachusetts, US	572 (3.5)	▲	230	92
2*	Minnesota, US	554 (5.9)	▲	257	106
2	Quebec, C	519 (3.0)	▲	221	93
2	Ontario, C	512 (3.1)	▲	226	90
2	Alberta, C	505 (3.0)	▲	217	88
2	British Columbia,C	505 (2.7)	▲	236	93
▶ **	Dubai, UAE	444 (2.1)	▼	296	123

Percentiles of performance



95% Confidence Interval for Average (±1.96 se)

▲ *Country average significantly higher than New Zealand average*

▲ *Country average significantly lower than New Zealand average*

Note: * *Met guidelines for sample participation rates only after replacement schools were included.*
** *Nearly satisfied guidelines for sample participation rates only after replacement schools were included.*
1 National Target Population does not include all of the International Target Population defined by TIMSS.
2 National Defined Population covers 90% to 95% of National Target Population.
▶ *Kuwait and Dubai, UAE tested the same cohort of students as other countries, but later in 2007, at the beginning of the next school year.*
Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Source: *Adapted from Exhibits 1.1 and D.1 Mullis, Martin, and Foy, 2008.*

Table 6 Selected contextual factors for TIMSS 2006/07 countries

Country	Years of formal schooling*	Average age at time of testing	Human Development Index**	Gross National Income per capita (in US dollars)***	GNI per capita (purchasing power parity)	Average hours of instructional time in mathematics (teacher reports)
Hong Kong SAR	4	10.2	0.937	29040	39200	150 (3.4)
Singapore	4	10.4	0.922	28730	43300	201 (0.8)
Chinese Taipei	4	10.2	0.932	17294	-	112 (2.6)
Japan	4	10.5	0.953	38630	32840	136 (1.2)
Kazakhstan	4	10.6	0.794	3870	8700	133 (1.8)
Russian Federation	4	10.8	0.813	5770	12740	110 (1.3)
England	5	10.2	0.946	40560	33650	183 (2.1)
Latvia	4	11.0	0.855	8100	14840	121 (3.1)
Netherlands	4	10.2	0.953	43050	37940	179 (4.6)
Lithuania	4	10.8	0.862	7930	14550	118 (1.7)
United States	4	10.3	0.951	44710	44070	171 (3.7)
Germany	4	10.4	0.935	36810	32680	145 (1.5)
Denmark	4	11.0	0.949	52110	36190	125 (1.2)
Australia	4	9.9	0.962	35860	33940	174 (5.4)
Hungary	4	10.7	0.874	10870	16970	110 (1.3)
Italy	4	9.8	0.941	31990	28970	201 (2.8)
Austria	4	10.3	0.948	39750	36040	126 (1.1)
Sweden	4	10.8	0.956	43530	34310	104 (2.3)
Slovenia	4	9.8	0.917	18660	23970	141 (1.0)
Armenia	4	10.6	0.775	1920	4950	133 (3.4)
Slovak Republic	4	10.4	0.863	9610	17060	143 (0.6)
Scotland	5	9.8	0.946	40560	33650	181 (2.7)
New Zealand	**4.5 - 5.5**	**10.0**	**0.943**	**26750**	**25750**	**148 (1.8)**
Czech Republic	4	10.3	0.891	12790	20920	144 (1.1)
Norway	4	9.8	0.968	68440	50070	115 (2.5)
Ukraine	4	10.3	0.788	1940	6110	104 (1.4)
Georgia	4	10.1	0.754	1580	3880	130 (1.5)
Iran, Islamic Rep. of	4	10.2	0.759	2930	9800	105 (2.6)
Algeria	4	10.2	0.733	3030	5940	177 (4.7)
Colombia	4	10.4	0.791	3120	6130	175 (4.7)
Morocco	4	10.6	0.646	2160	3860	162 (2.5)
El Salvador	4	11.0	0.735	2680	5610	147 (2.6)
Tunisia	4	10.2	0.766	2970	6490	166 (1.6)
Kuwait	4	10.2	0.891	30630	48310	x x
Qatar	4	9.7	0.875	-	-	x x
Yemen	4	11.2	0.508	760	2090	134 (7.2)

Note: * *Represents years of schooling counting from the first year of primary schooling.*
** *Taken from United Nations Development Programme's Human Development Report 2007/2008. See Mullis, Martin, and Foy for details.*
*** *Data on GNI taken from the World Bank's 2008 World Development Indicators. Purchasing Power Parity adjusts the GNI to take account of comparative purchasing power between the country and the United States.*
Standard errors are presented in parentheses.

Source: Adapted from Exhibits 3, 1.1, and 5.2, Mullis, Martin, and Foy, 2008.

International trends in mathematics achievement at the middle primary level

There are several ways that trends since 1994 can be examined for the countries participating in TIMSS. The analyses presented here will include only those countries that have participated in all three international cycles, 1994/95, 2002/03, and 2006/07. Table 7 shows the change in mean mathematics scores since 1994/95, ordered so that those countries that have had the biggest positive change since the first cycle are at the top and those with the biggest negative change are at the bottom.

Table 7 Trends in middle primary school mean mathematics achievement in three cycles of TIMSS

Country	1994/95 to 2006/07 difference		2002/03 to 2006/07 difference	
England	57 (4.4)	▲	10 (4.7)	▲
Hong Kong SAR	50 (5.4)	▲	32 (4.8)	▲
Slovenia	40 (3.6)	▲	23 (3.2)	▲
Latvia	38 (5.1)	▲	4 (3.8)	
New Zealand	**23 (5.0)**	▲	**-3 (3.2)**	
Australia	22 (4.9)	▲	17 (5.3)	▲
Iran, Islamic Rep. of	15 (6.6)	▲	13 (5.7)	▲
United States	11 (3.8)	▲	11 (3.4)	▲
Singapore	9 (5.9)		5 (6.7)	
Japan	1 (2.8)		4 (2.6)	
Scotland	1 (4.7)		4 (3.9)	
Norway	-3 (4.1)		22 (3.5)	▼
Hungary	-12 (5.1)	▼	-19 (4.8)	▼
Netherlands	-14 (3.7)	▼	-5 (3.0)	

Note: ▲ *2006/07 score significantly higher.*
▼ *2006/07 score significantly lower.*
Standard errors are presented in parentheses.
Source: Adapted from Exhibit 1.3 Mullis, Martin, and Foy, 2008.

England is the country with the largest change over time in mean mathematics score. Hong Kong SAR, Slovenia, Latvia, New Zealand and Australia have also all had significant increases in mean mathematics achievement since 1994. Both England and Hong Kong have undergone curriculum reforms since the first implementation of TIMSS. In addition Latvia and Slovenia have both made significant changes in their education systems since 1994; summaries of these along with summaries of the changes in England and Hong Kong are presented in the following paragraphs.

England

The English National Curriculum was revised in 1999/2000 and the non-statutory National Numeracy Strategy was introduced in 1998 then formally implemented at primary school level in 1999, influencing the teaching of mathematics particularly at primary level. The English curriculum is structured in three age related key stages: key stage 1 is years 1 and 2 of primary school, ages 5 to 7; key stage 2 is years 3 to 6 of primary school, ages 7 to 11; and key stage 3 is years 7 to 9 of secondary school, ages 11 to 14. Students are assessed in English and mathematics at the end of each of these key stages; for the students at ages 11 and 14, their tests are externally scored and the school-by-school results published nationally (Ruddock, 2008).

Hong Kong

In Hong Kong, a new curriculum framework (including new syllabuses) was introduced in primary schools in 2002, with a move away from traditional rote learning to a more student centred learning system. The Basic Competency Assessment (BCA), comprising a student assessment and a system assessment, was also introduced from 2003 to help monitor learning at key stages for Chinese, English and Mathematics (Grades 3, 6, and 9). The student assessment consists of an online resource bank that teachers can use to design appropriate assessment tasks for their students. The system assessment is administered at the territory level by the government and provides feedback to schools, which can then feed into schools' planning around effectiveness in learning and teaching (Tse & Loh, 2007 and Lam, 2002).

The Hong Kong government has also increased its focus on teacher education and qualifications. Mathematics specialists can be found teaching students at primary school level. From the 2004-2005 school year, all graduates of pre-service primary and secondary teacher education programmes have been degree holders. The percentage of primary school teachers who are degree holders has increased to 80.4 percent in the 2006-2007 school year, compared to 49.6 percent in the 2001-2002 school year (Leung & Leung, 2008).

Latvia

Since 1998, Latvia has had a basic education standard for students in grades 1 to 9 (aged 7 to 16).[7] Subject standards, which are part of the basic education standard, determine the main aims and tasks of the subject, the mandatory content of the subject, and the forms and order of the evaluation of achievement. The number of lessons per week is set nationally and mandatory. In grades 1 to 4 students have one teacher for all subjects; from grade 5, students have specialist subject teachers. Latvian students have tests in all grades, but the first national assessments occur at the end of grade 3 (students aged 9 and 10 – see Geske, Grinfelds, & Ozola, 2008).

Slovenia

Slovenia has been undergoing some significant changes in its schooling system, the most obvious of which is the lowering of the school starting age from 7 to 6 and revised national curricular documents for all levels of pre-university education. The goal of the reforms, implementation of which began in 1999, are:

"a higher level of interconnectedness of disciplinary knowledge, and increased active role of students, internationally comparable standards and levels of knowledge, improvement in functional literacy, and an increase in the quality and longevity of acquired knowledge." (Japelj Pavešić & Svetlik, 2008, p. 537).

The Slovenian syllabus specifies the exact number of yearly and weekly lessons for individual subjects. In grades 1 to 3, nearly all subjects are taught by general class teachers. During grades 4 to 6, specialist teachers become more and more involved in the teaching process.

Relative rankings

In many summaries of the international data, relative rankings of mean scores are used to describe change. This is not a particularly desirable practice as any mean scores derived from a sample and ascribed to a population have some level of uncertainty around them and rankings ignore this uncertainty. In addition, some presentations of rankings fail to mention the number of countries included in the ranking.

Table 8 presents relative ranking changes between 1994/95 and 2006/07. This should be read with caution, because, although a country may be ranked higher, the mean scores may not be significantly different when the uncertainties are taken into account. For example, the mean mathematics achievement for Hong Kong SAR and that of Singapore in 2006/07 are not significantly different.

Table 8 shows that for New Zealand, the mean mathematics scores have improved quite considerably between 1994 and 2006, and there has been very little movement in our position in the 'league tables' when only those countries in both assessments are included. Despite New Zealand's Year 5 mean mathematics score rising over time, it has remained lower than the mean of the 14 countries as this mean has also risen over time.

7 Pre-primary education is compulsory for students aged 5 and 6 in Latvia.

Table 8 **Middle primary mean mathematics scores for countries participating in three cycles of TIMSS from 1994/95 to 2006/07**

1994/95 mean mathematics score			2002/03 mean mathematics score			2006/07 mean mathematics score		
Singapore	590 (4.5)	▲	Singapore	594 (5.6)	▲	Hong Kong SAR	607 (3.6)	▲
Japan	567 (1.9)	▲	Hong Kong SAR	575 (3.2)	▲	Singapore	599 (3.7)	▲
Hong Kong SAR	557 (4.0)	▲	Japan	565 (1.6)	▲	Japan	568 (2.1)	▲
Netherlands	549 (3.0)	▲	Netherlands	540 (2.1)	▲	England	541 (2.9)	▲
Hungary	521 (3.6)	▲	Latvia	533 (3.1)	▲	Latvia	537 (2.3)	▲
United States	518 (2.9)	▲	England	531 (3.7)	▲	Netherlands	535 (2.1)	▲
Latvia	499 (4.6)		Hungary	529 (3.1)	▲	United States	529 (2.4)	▲
Australia	495 (3.4)	▼	United States	518 (2.4)		Australia	516 (3.5)	
Scotland	493 (4.2)	▼	Australia	499 (3.9)	▼	Hungary	510 (3.5)	▼
England	484 (3.3)	▼	**New Zealand**	**496 (2.1)**	▼	Slovenia	502 (1.8)	▼
Norway	476 (3.0)	▼	Scotland	490 (3.3)	▼	Scotland	494 (2.2)	▼
New Zealand	**469 (4.4)**	▼	Slovenia	479 (2.6)	▼	**New Zealand**	**492 (2.3)**	▼
Slovenia	462 (3.1)	▼	Norway	451 (2.3)	▼	Norway	473 (2.5)	▼
Iran, Islamic Rep. of	387 (5.0)	▼	Iran, Islamic Rep. of	389 (4.2)	▼	Iran, Islamic Rep. of	402 (4.1)	▼
Mean for all 14*	**505 (1.1)**		**Mean for all 14***	**514 (1.0)**		**Mean for all 14***	**522 (0.9)**	

Note: ▲ *Country mean is significantly higher than the mean for the 14 countries*
▼ *Country mean is significantly lower than the mean for the 14 countries*
** This mean has been calculated for the 14 countries common to all cycles. It is calculated by pooling all student results for the 14 countries and weighting so that each country contributes equally to the mean.*
Standard errors are presented in parentheses.

International trends in mathematics benchmarks

As shown in Table 9, five percent of New Zealand Year 5 students reached the advanced benchmark, the point at where students were deemed capable of applying *their understanding and knowledge in a variety of relatively complex situations and explain their reasoning*. This was a similar proportion to countries including Italy (6%), Germany (6%), the Slovak Republic (5%), and Scotland (4%) and higher than Slovenia (3%), Austria (3%), and Sweden (3%). However, Singapore was the country with the greatest proportion of students at the advanced benchmark, more than eight times the proportion of New Zealand students, at 41 percent.

Examining the low benchmark, 15 percent of New Zealand students did not reach this benchmark and therefore, in terms of the benchmark definition, did demonstrate *some basic mathematical knowledge*. Most countries had some students in this group, with Chinese Taipei (1%), Singapore (2%), Japan (2%) and the Netherlands (2%) having the fewest students unable to reach the low benchmark. All students in Hong Kong SAR were at least able to reach this low benchmark. Countries with similar proportions to New Zealand at the advanced benchmark generally had fewer students unable to reach the low benchmark when compared to New Zealand.

Included in the table is the international median percentage of students at each benchmark. The same proportion of New Zealand Year 5 students reached the advanced and high benchmarks as the international median, so New Zealand was around the middle of the countries for these benchmarks. For the intermediate and low benchmarks, proportionally fewer New Zealand Year 5 students reached these benchmarks compared to the international median.

Table 9 Proportion of middle primary students at each international benchmark

	Country	Percentage of students at each benchmark: Advanced	High	Intermediate	Low
	Singapore	41 (2.1)	74 (1.7)	92 (0.9)	98 (0.3)
	Hong Kong SAR	40 (2.2)	81 (1.6)	97 (0.5)	100 (0.1)
	Chinese Taipei	24 (1.2)	66 (1.2)	92 (0.5)	99 (0.2)
	Japan	23 (1.2)	61 (1.2)	89 (0.8)	98 (0.4)
1	Kazakhstan	19 (2.1)	52 (3.5)	81 (2.9)	95 (1.5)
	England	16 (1.2)	48 (1.4)	79 (1.2)	94 (0.7)
	Russian Federation	16 (1.8)	48 (2.3)	81 (1.7)	95 (0.7)
1	Latvia	11 (0.8)	44 (1.5)	81 (1.2)	97 (0.5)
2 *	United States	10 (0.8)	40 (1.3)	77 (1.2)	95 (0.5)
1	Lithuania	10 (0.7)	42 (1.4)	77 (1.4)	94 (0.7)
	Hungary	9 (0.8)	35 (1.4)	67 (1.7)	88 (1.2)
	Australia	9 (0.8)	35 (1.9)	71 (1.7)	91 (1.0)
	Armenia	8 (1.5)	28 (1.8)	60 (1.8)	87 (1.2)
*	Denmark	7 (0.7)	36 (1.5)	76 (1.2)	95 (0.8)
**	Netherlands	7 (0.7)	42 (1.6)	84 (1.3)	98 (0.4)
	Germany	6 (0.5)	37 (1.3)	78 (1.2)	96 (0.5)
	Italy	6 (0.7)	29 (1.6)	67 (1.6)	91 (1.0)
	New Zealand	**5 (0.5)**	**26 (1.0)**	**61 (1.1)**	**85 (1.0)**
	Slovak Republic	5 (0.7)	26 (1.4)	63 (1.8)	88 (1.5)
*	Scotland	4 (0.5)	25 (1.1)	62 (1.4)	88 (0.9)
	Slovenia	3 (0.4)	25 (1.1)	67 (0.9)	92 (0.6)
	Austria	3 (0.3)	26 (1.0)	69 (1.4)	93 (0.8)
	Sweden	3 (0.3)	24 (1.4)	68 (1.4)	93 (0.7)
	Ukraine	2 (0.5)	17 (1.1)	50 (1.5)	79 (1.2)
	Czech Republic	2 (0.4)	19 (1.4)	59 (1.6)	88 (1.1)
	Norway	2 (0.3)	15 (1.0)	52 (1.6)	83 (1.1)
1	Georgia	1 (0.4)	10 (1.0)	35 (1.8)	67 (2.0)
	Colombia	0 (0.1)	2 (0.4)	9 (1.1)	31 (2.0)
	Morocco	0 (0.2)	2 (0.8)	9 (1.1)	26 (2.0)
	Iran, Islamic Rep. of	0 (0.1)	3 (0.5)	20 (1.5)	53 (2.0)
	Algeria	0 (0.1)	2 (0.4)	14 (1.4)	41 (2.2)
	Tunisia	0 (0.1)	1 (0.2)	9 (0.7)	28 (1.6)
	El Salvador	0 (0.0)	1 (0.2)	6 (0.5)	22 (1.6)
▶	Kuwait	0 (0.0)	0 (0.1)	5 (0.6)	21 (1.2)
	Qatar	0 (0.0)	0 (0.1)	2 (0.2)	13 (0.4)
	Yemen	0 (0.0)	0 (0.1)	1 (0.4)	6 (0.8)
International Median		**5**	**26**	**67**	**90**
Benchmarking Participants					
2	Massachusetts, US	22 (1.8)	63 (2.1)	92 (1.1)	99 (0.3)
2 *	Minnesota, US	18 (2.1)	55 (3.2)	85 (2.2)	97 (1.2)
2	Quebec, C	5 (0.7)	34 (2.2)	74 (1.6)	96 (0.6)
2	British Columbia, C	4 (0.5)	27 (1.3)	67 (1.7)	93 (0.9)
2	Ontario, C	4 (0.6)	29 (1.8)	71 (1.8)	94 (1.1)
2	Alberta, C	3 (0.6)	25 (1.8)	69 (1.9)	94 (1.0)
▶ **	Dubai, UAE	2 (0.3)	12 (0.7)	37 (1.2)	69 (1.3)

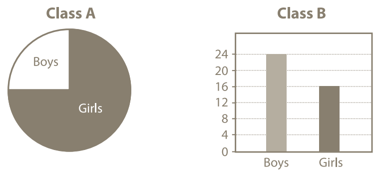
Note: * *Met guidelines for sample participation rates only after replacement schools were included.*
** *Nearly satisfied guidelines for sample participation rates only after replacement schools were included.*
1 *National Target Population does not include all of the International Target Population defined by TIMSS.*
2 *National Defined Population covers 90% to 95% of National Target Population.*
▶ *Kuwait and Dubai, UAE tested the same cohort of students as other countries, but later in 2007, at the beginning of the next school year.*
Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Figures 3 to 6 present examples of questions that Year 5 students achieving at or above the *advanced, high, intermediate*, and *low* benchmarks were likely to have answered correctly. An example of a correct answer and a summary of the scoring guide are presented. In addition, proportions of students successfully completing the question for a selection of countries, including the best performing country on that question, are shown. The international average is also presented as an indication of how students in all 37 countries performed on this question.

Figure 3 Question students reaching the advanced benchmark are likely to have answered correctly

Question with example of correct answer
Content domain: data display
Cognitive domain: reasoning

Class A and B each have 40 students



There are more girls in Class A than in Class B.
How many more?

- ● 14
- Ⓑ 16
- Ⓒ 24
- Ⓓ 30

Scoring guide

Uses data from two different graph types to solve a problem.

Country	Percent full credit	
Singapore	63	(2.3)
Hong Kong SAR	63	(2.3)
Chinese Taipei	47	(2.5)
Russian Federation	42	(3.0)
Japan	41	(2.2)
England	40	(2.5)
United States	38	(1.8)
Australia	36	(2.2)
Scotland	34	(2.3)
International Avg.	32	(0.4)
New Zealand	**32**	**(1.6)**

Note: Standard errors are presented in parentheses.
Source: Adapted from Exhibit 2.6, Mullis, Martin, and Foy, 2008.

Figure 4 Question students reaching the high benchmark are likely to have answered correctly

Question with example of correct answer
Content domain: number
Cognitive domain: reasoning

Mano did the subtraction problem above for homework but spilled some of his drink on it. One digit could not be read. His answer of 415 was correct. What is the missing digit?

Answer: 2

Scoring guide

Determines the missing digit to give a specified difference in a three-digit subtraction problem.

Country	Percent full credit	
Chinese Taipei	88	(1.6)
Hong Kong SAR	85	(1.9)
Singapore	85	(1.4)
Russian Federation	84	(1.8)
Japan	80	(1.8)
International Avg.	42	(0.4)
United States	41	(1.8)
England	28	(2.1)
Scotland	25	(2.2)
Australia	22	(2.6)
New Zealand	**18**	**(1.6)**

Note: Standard errors are presented in parentheses.
Source: Adapted from Exhibit 2.8, Mullis, Martin, and Foy, 2008.

Figure 5 Question students reaching the intermediate benchmark are likely to have answered correctly

Question with example of correct answer
Content domain: number
Cognitive domain: applying

Al wanted to find how much his cat weighed. He weighed himself and noted that the scale read 57kg. He then stepped on the scale holding his cat and found that it read 62kg.

What was the weight of the cat in kilograms?

Answer: 5 kilograms

Scoring guide

Solves a measurement word problem involving subtraction of two-digit numbers.

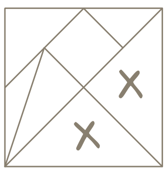
Country	Percent full credit	
Chinese Taipei	95	(1.2)
Singapore	87	(1.3)
Russian Federation	86	(1.8)
Hong Kong SAR	86	(1.7)
Japan	83	(2.0)
Scotland	64	(2.7)
England	63	(2.2)
Australia	61	(2.4)
International Avg.	60	(0.3)
United States	60	(1.7)
New Zealand	**53**	**(2.1)**

Note: Standard errors are presented in parentheses.
Source: Adapted from Exhibit 2.12, Mullis, Martin, and Foy, 2008.

Figure 6 Question students reaching the low benchmark are likely to have answered correctly

Question with example of correct answer
Content domain: geometric shapes and measures
Cognitive domain: knowing

The square is cut into 7 pieces. Put an X on each of the 2 triangles that are the same size and shape.



Scoring guide:

Identifies two triangles with the same size and shape in a complex figure.

Country	Percent full credit	
Hong Kong SAR	91	(1.2)
Scotland	88	(1.4)
England	88	(1.4)
Singapore	88	(1.4)
Japan	87	(1.4)
Australia	85	(1.9)
United States	85	(1.0)
Russian Federation	81	(2.6)
Chinese Taipei	81	(1.9)
New Zealand	**81**	**(1.4)**
International Avg.	72	(0.3)

Note: Standard errors are presented in parentheses.
Source: Adapted from Exhibit 2.15, Mullis, Martin, and Foy, 2008.

International trends in mathematics content and cognitive domains

As mentioned earlier, New Zealand Year 5 students demonstrated a relative strength in data display questions and a relative weakness in number questions. Relatively higher data display mean scores and relatively lower number mean scores were also observed for Japan and Scotland. In contrast, the higher achieving countries, Singapore, Chinese Taipei, and Hong Kong SAR all showed a relative strength in the number domain.

In the cognitive domains, New Zealand Year 5 students demonstrated a relative strength in questions that required them to use their reasoning and were relatively worse at questions that required demonstrating their knowledge. None of the other English-speaking or high-performing countries had a pattern quite like this. Scotland's and Australia's students demonstrated a relative weakness in the knowing domain and a relative strength in the applying domain. In contrast, Chinese Taipei, Hong Kong SAR, Singapore, England and the United States all showed a relative strength in the knowing domain.

TIMSS and the New Zealand mathematics curriculum

Mathematics curriculum levels and the TIMSS content domains

In order to gain greater understanding of the relationship between the *Mathematics in the New Zealand Curriculum* (MiNZC)[8] levels and student achievement in TIMSS, New Zealand teachers were asked at which level of the MiNZC most of the students in their class were currently working for each of the strands: *Number, Algebra, Measurement, Geometry, and Statistics*.[9] Note that the information was not collected for individual students but for the majority of the TIMSS students the teacher taught. For the purpose of analysis, a teacher's response has been assigned to each individual student in the class. Figure 7 shows that while the majority of Year 5 students were working at level 3 of the curriculum, there were still a significant number of students in classes working at level 2, particularly in the Algebra and Measurement strands.

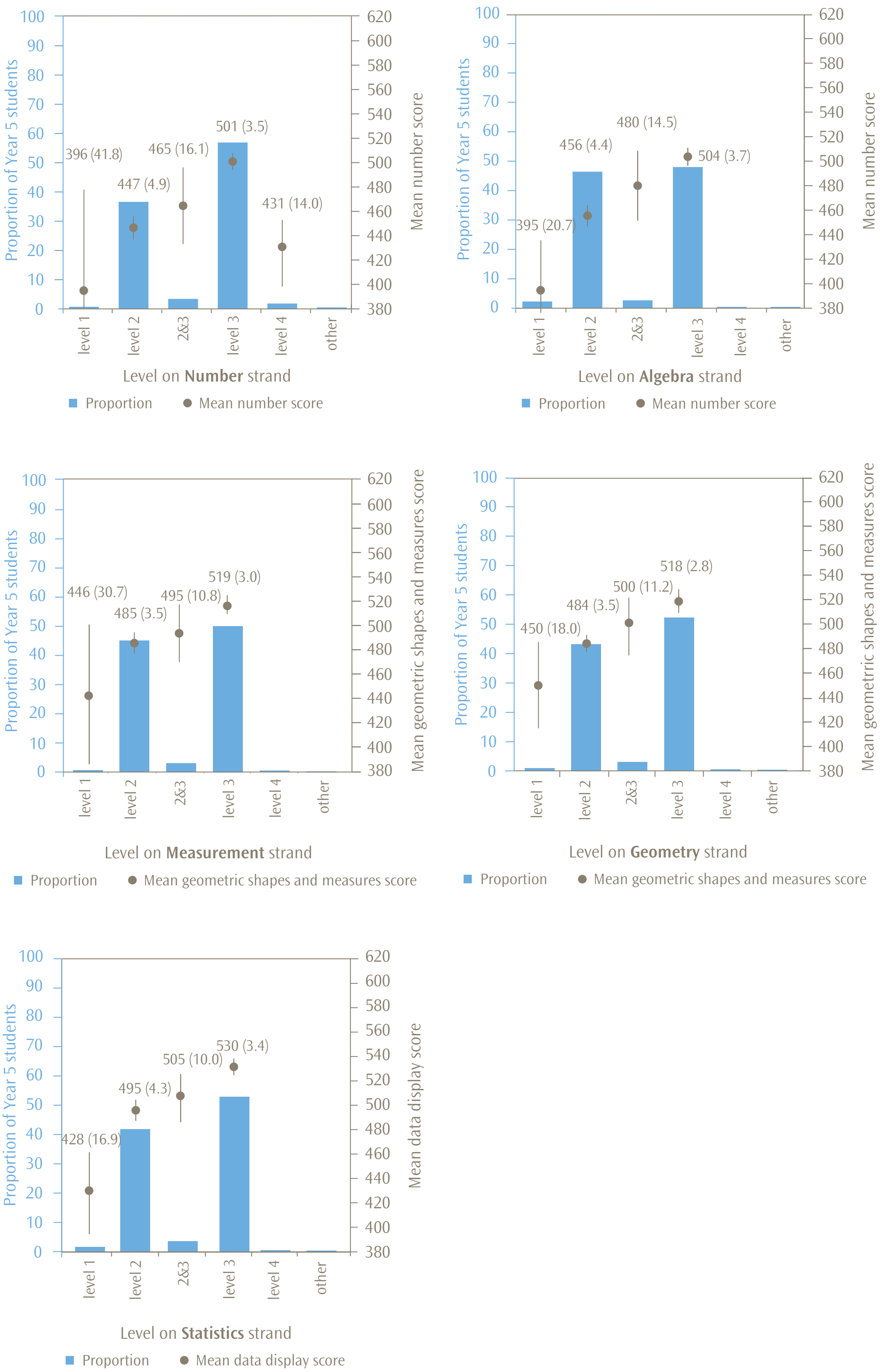
Since the TIMSS domains were similar to the MiNZC strands in terms of content, with *Measurement* and *Geometry* combined similar to the *geometric shapes and measures* domain, and *Number* and *Algebra* combined similar to the *number* domain, the figure also maps the mean TIMSS domain score for the students estimated to be working at each level of the MiNZC. For example, the mean score in the TIMSS *number* domain for the 57 percent of students in classes estimated to be working at level 3 of the *Number* strand of the MiNZC was 501 scale score points. The figure shows that students whose classes are working at higher levels of the curriculum have higher achievement on the associated TIMSS content domain. Note that no attempt is being made here to infer a causal link – that is, we are not saying the higher mean achievement is **because** they are working at the higher level.

It is interesting to look at these results in an international context and observe that if only those students working at level 3 of the curriculum were included in the TIMSS testing, New Zealand's overall mathematics score would still have been below that of the high-performing countries, Singapore, Chinese Taipei, Hong Kong SAR, and Japan. For example, the mean score for Singaporean students on the number domain was 611 scale score points, while New Zealand students working at level 3 had a mean score of 501 scale score points.

8 This was the curriculum in place at the time of testing.

9 The Mathematical Processes strand of the MiNZC relates more to the TIMSS cognitive domains and so is not included in this discussion

Figure 7 Mean content area achievement by New Zealand curriculum strands



Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. In the cases where there are no mean scores, the 'other' grouping, there were too few students to report achievement. Scores presented for levels 1, 4, and combined 2 & 3 should be treated with caution as the proportion of students is in each of these groups is small. Standard errors are presented in parentheses.

Curriculum match

Questions about international studies often focus on the appropriateness of the assessment questions for New Zealand students. New Zealand is not unique in asking this question; other countries are also concerned with appropriateness of the tests. The TIMSS assessment questions are developed through a collaborative process that begins with the development of an assessment framework. The *TIMSS 2007 assessment frameworks* (Mullis, Martin, et al., 2005) were designed to specify the important aspects of mathematics that participating countries agreed should be the focus of an international assessment of mathematics achievement. However it is inevitable that the tests included questions that were unfamiliar to some students in some countries. In order to investigate the extent to which the TIMSS 2006/07 assessment was relevant to each country's curriculum, TIMSS conducted a Test-Curriculum Matching Analysis (TCMA). The TCMA was also used to investigate the impact of selecting only appropriate questions on a country's performance.

For the TCMA, each question was examined using the following two criteria:

- whether or not the topic of the question is in the intended curriculum for the majority of middle primary students (in our case Year 5) – that is, more than 50 percent; and
- whether or not the question topic is intended to be encountered by the middle primary students prior to the TIMSS testing (testing for New Zealand Year 5 students occurred in the beginning of November).

While all questions, regardless of this analysis, were included in any overall results reported for TIMSS, this analysis was used to ascertain the level to which the results might change for New Zealand if only questions judged appropriate were included in the tests. The analysis also included an examination of how students in other countries would fare if given only the "New Zealand-appropriate" test.

Table 10 shows the proportion of questions considered appropriate to the New Zealand curriculum in each of the TIMSS content areas. However, it should be noted that New Zealand's mathematics curriculum provides some challenges for deciding whether at least half of Year 5 students are likely to have met the question topics in the TIMSS test.[10] The curriculum is not prescriptive, instead providing some broad guidelines of mathematics concepts and skills that schools can choose to cover. Schools are encouraged to design mathematics programmes that are relevant to their students and communities. Consequently, when schools plan their mathematics programmes there is considerable variation between them. Another challenge is that the broad achievement objectives are grouped in levels which cover approximately two years of schooling. As shown in the previous section, New Zealand Year 5 students were generally working at levels 2 and 3 of the curriculum, so information from levels 1, 2, and 3 was used to guide judgements on the TCMA.[11]

Table 10 Appropriateness of the TIMSS tests to the New Zealand curriculum

TIMSS content domain	Number of score points judged appropriate for New Zealand curriculum	Number of score points in TIMSS assessment	Proportion of score points judged appropriate for New Zealand curriculum
Number	82	98	84%
Geometric shapes and measures	59	65	91%
Data display	24	28	86%

Note: Number corresponds to the Number and Algebra strands of the curriculum, geometric shapes and measures corresponds to the Measurement and Geometry strands of the curriculum, and data display corresponds to the Statistics strand of the curriculum.

Although only around 85 percent of the questions were judged appropriate for New Zealand students, the TIMSS TCMA analysis shows that some of the higher-performing countries would have done better on the 'New Zealand' test than New Zealand Year 5 students, as shown in Table 11.

10 Note that for the TCMA, the curriculum document used was the 1992 Mathematics in the New Zealand Curriculum (Ministry of Education, 1992).

11 Thanks to Chris Joyce and Ally Bull from NZCER for this work.

Table 11 **Performance of middle primary students from selected countries on the 'New Zealand' appropriate test**

Country	Average percent correct on New Zealand test
Hong Kong SAR	77
Singapore	74
Chinese Taipei	70
Japan	68
Russian Federation	62
England	63
United States	60
Australia	57
Scotland	52
New Zealand	**51**

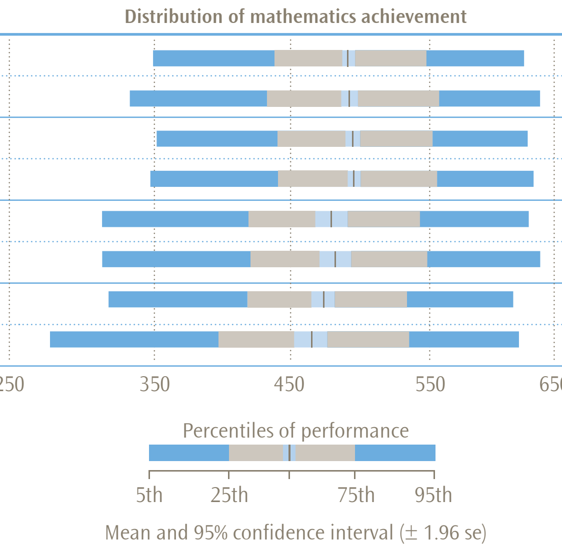
Source: Adapted from Exhibit C.1 in Mullis, Martin, and Foy, 2008.

Mathematics achievement by gender

There was no significant difference in mean mathematics achievement between Year 5 boys (493) and girls (492) in 2006. However, the distribution of achievement was wider for boys (297) than for girls (268) as shown in Figure 8. As Figure 8 also shows, the wider distribution among boys has been consistent over the four cycles. Similarly, there was no significant difference between the mean mathematics achievement of boys and that of girls in each of the preceding cycles. It is encouraging to observe the narrowing of the distribution of mathematics achievement for both boys and girls between 1994 and 2006, with fewer lower performers in 2006 compared with 1994.

Figure 8 Trends in distributions of achievement for girls and boys from 1994 to 2006

Year		Mean mathematics score	Range from 5th to 95th percentile	Inter-quartile range from 25th to 75th percentile
2006	Girls	492 (2.4)	268	110
	Boys	493 (3.1)	297	125
2002	Girls	495 (2.8)	268	112
	Boys	496 (2.4)	277	115
1998	Girls	480 (6.0)	309	124
	Boys	482 (5.8)	317	128
1994	Girls	474 (4.3)	293	116
	Boys	465 (6.1)	339	138



Note: Standard errors are presented in parentheses.

Benchmarks for Year 5 boys and girls

There is a statistically significant difference between the proportions of girls and boys reaching the advanced and high benchmarks. Fewer girls are reaching these top benchmarks. At the lower levels however, there is not much difference between the proportions of girls and boys reaching the intermediate and low benchmarks. Similar proportions of girls and boys did not reach the low benchmark (13% and 16% respectively). In terms of the benchmark definitions, these boys and girls who did not reach the low benchmark did not demonstrate some basic mathematical knowledge.

Table 12 Proportion of Year 5 students reaching each international benchmark by gender in TIMSS 2006/07

	Percentage of Year 5 students reaching each benchmark			
Gender	Advanced	High	Intermediate	Low
Girls	4 (0.5)	24 (1.3)	61 (1.7)	87 (1.1)
Boys	6 (0.8)	28 (1.3)	60 (1.6)	84 (1.4)

Note: Standard errors are presented in parentheses.

The proportions of boys and girls reaching each of the benchmarks in 2006 were very similar to 2002 (see Caygill, Sturrock, & Chamberlain, 2007, p. 42). However, in comparison to 1994, proportionally more boys and girls reached the high, intermediate and low benchmarks in 2006. In addition, significantly more boys reached the advanced benchmark in 2006 compared with 1994. For girls, there was no significant difference between the proportions reaching the advanced benchmark in 2006 compared with 1994. Proportionally fewer girls and boys were lower achievers, that is, did not reach the low benchmark, in 2006 compared with 1994

Achievement on the content and cognitive domains for girls and boys

While there were no overall differences in mean mathematics achievement between girls and boys, there were some distinct differences in terms of the content domains. On average, girls had higher scores in data display, while boys had higher scores in number. Boys and girls performed similarly on the geometric shapes and measures domain and over the cognitive domains.

Table 13 Year 5 mean mathematics scores on the content and cognitive domains by gender

	Mean domain score	
Content domain	Girls	Boys
Number	474 (2.9)	482 (3.3) ▲
Geometric Shapes & Measures	504 (2.7)	500 (2.8)
Data Display	517 (3.1) ▲	509 (3.1)

	Mean domain score	
Cognitive domain	Girls	Boys
Knowing	482 (2.8)	482 (3.1)
Applying	494 (2.7)	497 (2.7)
Reasoning	503 (3.2)	503 (3.2)

Note: ▲ mean domain score significantly higher than other gender.
Standard errors are presented in parentheses.
Source: Exhibit 3.3 Mullis, Martin, and Foy, 2008.

Mathematics achievement by ethnicity, language, and country of birth

This section will examine the mathematics achievement of students in TIMSS across different ethnic groups, by use of English at home, and by country of birth. These three characteristics of students are interrelated so in the final part of this section they are examined together. This section will examine relationships with mathematics achievement, but it should be noted that the existence of a relationship does not infer a causal link.

Mathematics achievement by ethnicity

Five broad ethnic classifications are used to describe ethnicity in New Zealand. They are: Pākehā/European, Māori, Pasifika, Asian, and 'Other' ethnic groupings. The majority of Year 5 students in New Zealand were identified by their schools[12] as Pākehā/European (61%) or Māori (19%). Pasifika (10%) and Asian (7%) students made up most of the rest of the ethnic groupings, with four percent of students categorised in the Other ethnic grouping.

Previous cycles of TIMSS have shown that average mathematics achievement varies across ethnic groups. Although the variation in achievement is not caused by ethnicity *per se*, education policies have been introduced so that all students may realise their potential. Specific areas of focus for the Ministry of Education include the achievement of Māori and Pasifika students (Ministry of Education 2007). The results at the Year 5 level in TIMSS 2002/03 (Caygill, Sturrock, & Chamberlain, 2007) showed an increase in mathematics performance, on average, for Māori and Pasifika students since the first cycle in 1994/1995.

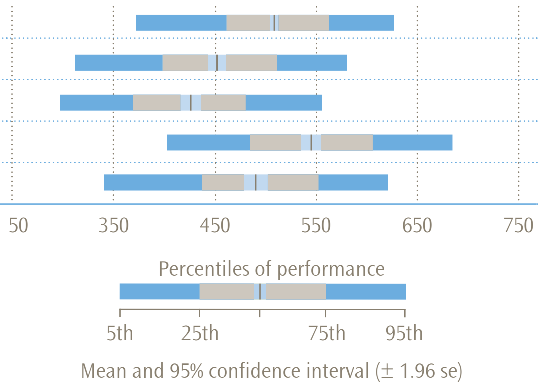
In TIMSS 2006/07, Asian (546) and Pākehā/European (510) students had significantly higher mean mathematics achievement than did their Māori (453), Pasifika (427) and Other (491) counterparts, as shown in Figure 9. On average, Asian students performed significantly higher in mathematics than Pākehā/European students and Māori students performed significantly higher than Pasifika students.

The distribution of achievement of Asian students and those in the Other ethnic grouping was the widest, while the distribution for the Pākehā/European students was the narrowest. Note that the 5th and 95th percentiles of achievement for the students in the Pasifika, Asian, and Other ethnic grouping should be treated with caution as there are few students at these ends of the distribution due to the smaller number of students in these ethnic groupings.

12 Based on enrolment information supplied by parents.

Figure 9 Distribution of New Zealand Year 5 mathematics achievement for each ethnic grouping

Ethnic grouping	Mean mathematics score	Distribution of mathematics achievement	Range from 5th to 95th percentile	Inter-quartile range from 25th to 75th percentile
Pākehā/European	510 (2.1)		255	101
Māori	453 (4.4)		269	113
Pasifika	427 (5.1)		259	112
Asian	546 (5.0)		282	122
Other	491 (6.0)		281	115



Note: The distribution and ranges for the students in the Pasifika, Asian, and Other ethnic groupings should be read with caution as there is a lot of uncertainty at the extremes of the distribution.
Standard errors are presented in parentheses.

Benchmarks for ethnic groupings

Within all ethnic groupings, there were students who reached the advanced benchmark; in terms of the benchmark definitions, they demonstrated the ability to complete tasks requiring applying their *understanding and knowledge in a variety of relatively complex situations and explain their reasoning*. Similarly, within all ethnic groupings there were students who did not reach the low benchmark; that is, they did not demonstrate the ability to complete a reasonable number of the simplest mathematics tasks which TIMSS seeks to measure.

Higher proportions of Asian and Pākehā/European students reached the advanced benchmark compared with each of the other ethnic groupings (as shown in Table 14). Around one-third of Pasifika students did not reach the low benchmark, while a quarter of Māori students did not reach this low benchmark.

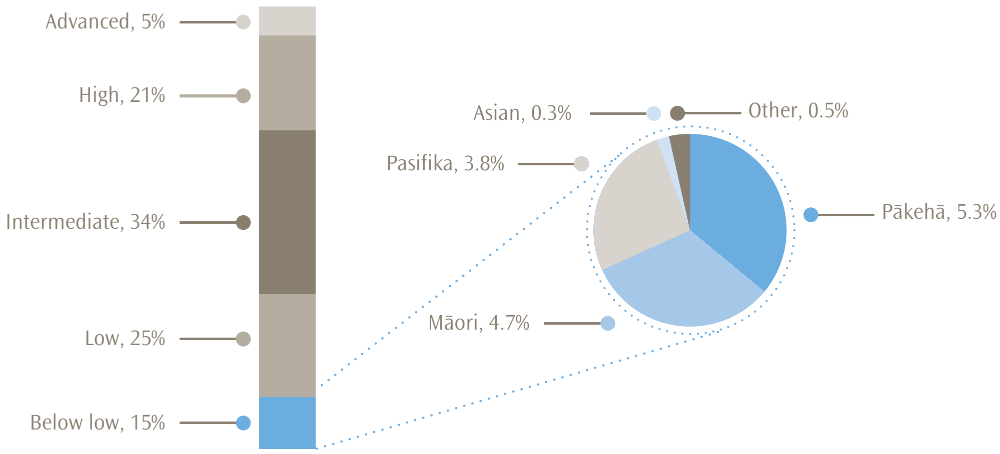
Table 14 Proportion of Year 5 students reaching each international benchmark by ethnic grouping

Ethnic grouping	Percentage of Year 5 students reaching each benchmark: Advanced	High	Intermediate	Low
Pākehā/European	5 (0.8)	32 (1.3)	70 (1.1)	91 (1.0)
Māori	1 (0.6)	12 (1.5)	41 (2.4)	75 (2.5)
Pasifika	1 (0.6)	7 (1.1)	28 (2.6)	62 (3.1)
Asian	19 (1.9)	50 (3.1)	78 (2.8)	95 (1.5)
Other	5 (1.8)	27 (3.8)	60 (3.4)	85 (3.4)

Note: Standard errors are presented in parentheses.

Another way of looking at this information is to examine the composition of the group who did not reach the low benchmark. Fifteen percent of New Zealand students did not reach this benchmark as shown in Figure 10. The majority of these students were Pākehā/European (5.3%) or Māori (4.7%). However, Māori and Pasifika students were over-represented in this lower-achieving group compared to their proportions in the population.

Figure 10 Ethnic composition of the students who did not reach the low benchmark



Note: The values presented in the pie chart are proportions of the whole population and therefore add to 15%, the proportion of students in the below low' group.

Trends in mean mathematics achievement for ethnic groupings

Pākehā/European, Māori, Pasifika, and Other students all demonstrated significant gains in mathematics achievement, on average, between 1994 and 2002. Asian students showed a positive change over the eight years between 1994 and 2002 but because of the large standard errors involved, this is not statistically significant. While between 2002 and 2006, the average performance of Māori students decreased significantly, the average mathematics achievement of Māori students in 2006 was significantly higher than in 1994. Between 2002 and 2006, the average performance of Pasifika students returned to the lower level of achievement observed in 1994. Asian students continued to show significant gains between 2002 and 2006, making a total gain of 62 scale points for Asian students between 1994 and 2006, a greater increase than for the remaining ethnic groups (see Table 15).

Table 15 Trends in mathematics achievement over four cycles of TIMSS by ethnic grouping

	Mean mathematics achievement								Change	
Ethnic grouping	1994		1998		2002		2006		1994 to 2006	
Pākehā/European	493	(3.9)	502	(5.0)	506	(2.7)	510	(2.1)	17	(4.4)
Māori	427	(8.2)	445	(7.3)	479	(4.8)	453	(4.4)	25	(9.3)
Pasifika	412	(11.0)	416	(15.1)	464	(6.3)	427	(5.1)	15	(12.2)
Asian	483	(16.9)	516	(9.9)	500	(6.0)	546	(4.9)	62	(17.7)
Other	475	(15.1)	481	(14.8)	504	(9.8)	491	(6.0)	16	(16.3)

Note: Standard errors are presented in parentheses.
Due to rounding some results may appear inconsistent.

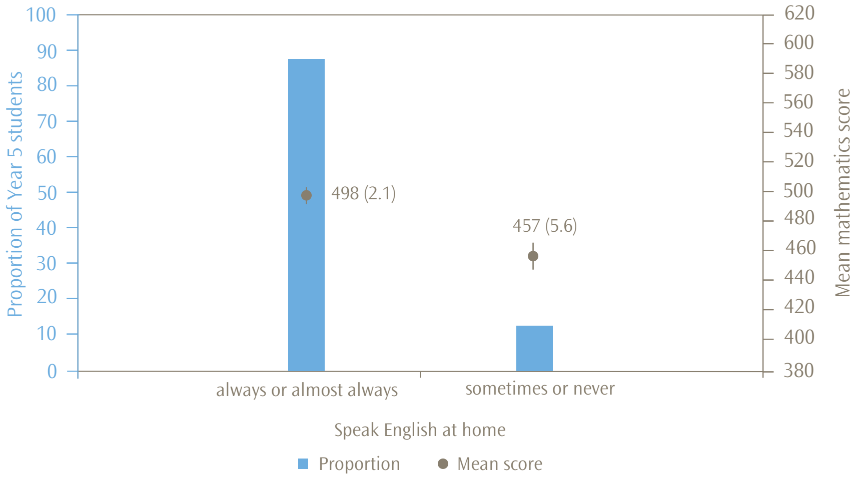
Mathematics achievement of boys and girls within ethnic groups

As mentioned earlier, there was no significant difference in mean mathematics achievement between boys and girls in TIMSS 2006/07. This result was also observed when gender differences were examined within each of the ethnic groups.

Mathematics achievement by regularity of English speaking at home

Most students reported that they always or almost always spoke the language of the test (in this case English) at home (87% - 74% always and 13% almost always).[13] Few students (1%) reported that they never spoke English at home. Students who always or almost always spoke English at home had higher mathematics achievement, on average, than those who sometimes or never spoke English at home (see Figure 11). This pattern of higher average achievement for those who spoke English at home was also evident across the previous three cycles of TIMSS (see Caygill, Sturrock, & Chamberlain, 2007). However, it is interesting to note that the difference between these two groups of students has reduced over time from 77 scale score points in 1994 to 41 scale score points in 2006.

Figure 11 Year 5 mean mathematics scores by regularity of English speaking at home



Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.

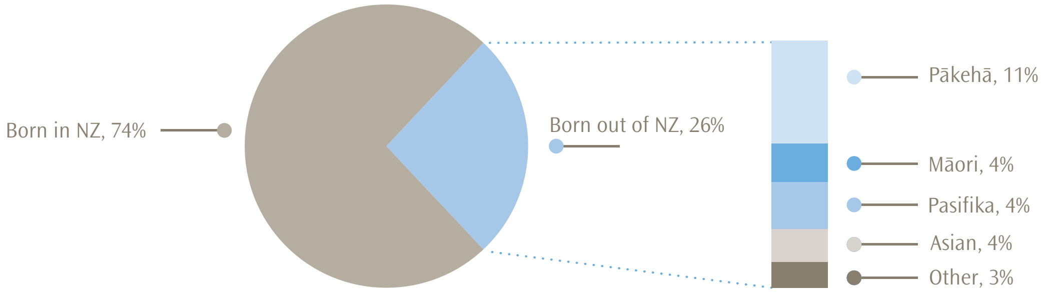
Mathematics achievement by country of birth

Another factor that interacts with language and ethnicity is the immigrant status of the student and their parents. This information was collected in TIMSS by asking the student if they and their parents were born in New Zealand. Around one-fifth of students had neither parent born in New Zealand, one-fifth had only one parent born in New Zealand and the rest had both parents born in New Zealand. One-quarter of students were born outside of New Zealand. Of these students born outside of New Zealand, nearly half of them (44%) came to New Zealand as school-age children.

Mathematics achievement was lower for those students born outside of New Zealand, on average, compared with the New Zealand born students (38 scale score points difference). The majority of the students born outside of New Zealand were Pākehā/European in ethnic origin, as shown in Figure 12.

13 In TIMSS 2006, as in 1994 and 1998, students who had the majority of their teaching in te reo Māori were excluded from the assessment. See technical notes and definitions for further details of exclusions.

Figure 12 Proportions of students born out of New Zealand by ethnic grouping



Note: The values presented in the bar are proportions of the whole population and therefore add to 26%, the proportion of students in the 'born out of NZ' group.

Interaction of use of English at home, ethnicity and country of birth

In order to confirm the relationships between use of English at home, ethnicity, country of birth and mathematics achievement and also to see how they interacted together, multiple-regression techniques were used. Only these background characteristics were included in the investigation. The resulting statistical model showed that speaking English infrequently at home, belonging to the Māori or Pasifika ethnic grouping, and being born outside of New Zealand, were all associated with lower mathematics achievement when the other factors were taken into account.[14] The model also demonstrates that all of these background characteristics were significant when explaining differences in mathematics achievement. Note that differences in achievement were smaller when the other factors were taken into account. For example, the difference between those who regularly spoke English at home and those who did not was reduced from 41 scale score points when analysed in isolation to 27 scale score points in the model. However, there are a limited number of factors included in this model. Taking into account socio-economic or educational resource factors may change this result (see the section later in this report entitled *Discussion of interactions*).

14 The model showed that when the other factors were taken into account, speaking English infrequently (-27 scale score points or ssp), Māori (-52 ssp), Pasifika (-62 ssp), born outside New Zealand (-35 ssp) were all associated with lower achievement.

Mathematics achievement by socio-economic status and home educational resources

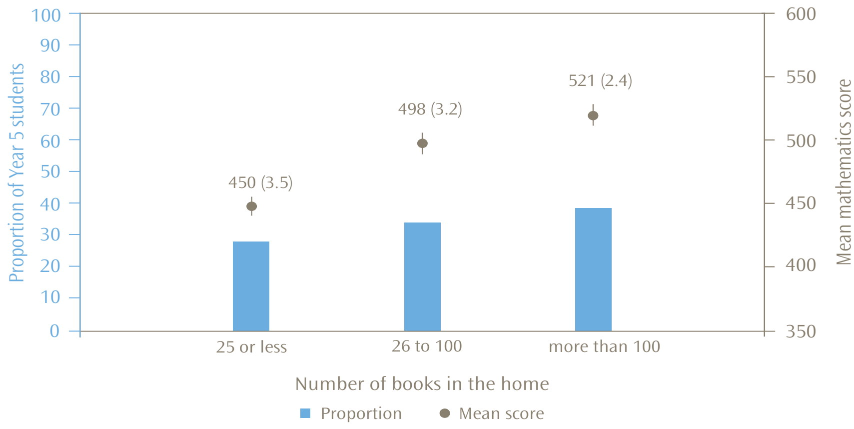
TIMSS includes a number of questions about resources available in the home. These resources in the home can be used as a proxy measure for socio-economic status. In addition, in New Zealand the decile indicator of schools is available to give a measure of the socio-economic status of the area in which a student lives. This section will present analyses of these proxy measures of socio-economic status and their association with mathematics achievement.

Number of books in the home

Just over one third of New Zealand Year 5 students (38%) reported having more than 100 books in their homes in 2006. This was a large reduction in proportion since 1994 when 62 percent reported having more than 100 books in their homes, but is consistent with 2002 and with that previously found by Caygill and Chamberlain (2005) in the 2001 Trends in Reading Literacy Study (38% also). This trend of fewer books in the home is also consistent with other countries that have been in the study since 1994. Thirty-four percent of students reported having between 25 and 100 books in their homes while 28 percent of students reported having 25 or fewer books in their homes.

As shown in Figure 13, there was a positive relationship between the number of books in the home and achievement in 2006, with those students with a greater number of books in the home having higher achievement, on average, in mathematics. This is consistent with findings from previous cycles of TIMSS as well as other studies that have shown a strong link between books in the home and achievement (see for example Chamberlain 2008).

Figure 13 **Proportions and mean mathematics achievement of Year 5 students by number of books in the home**



Note: *The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.*

Number of items in the home including educational resources

Students were asked whether their home contained items from a list of nine: *calculator, computer (do not include PlayStation®, GameCube®, XBox®, or other TV/video game computers), study desk/table for your use, dictionary, internet connection, your own room, your own mobile phone, musical instruments (e.g., piano, violin, guitar),* and *dishwasher*. The intention of this question was two-fold. The first four items were included to give an indication of the availability of resources at home that could be used to help educationally. The list in its entirety was included to give a proxy measure of socio-economic status as the students were too young to give reliable information on parental employment or household income.

Items in the home

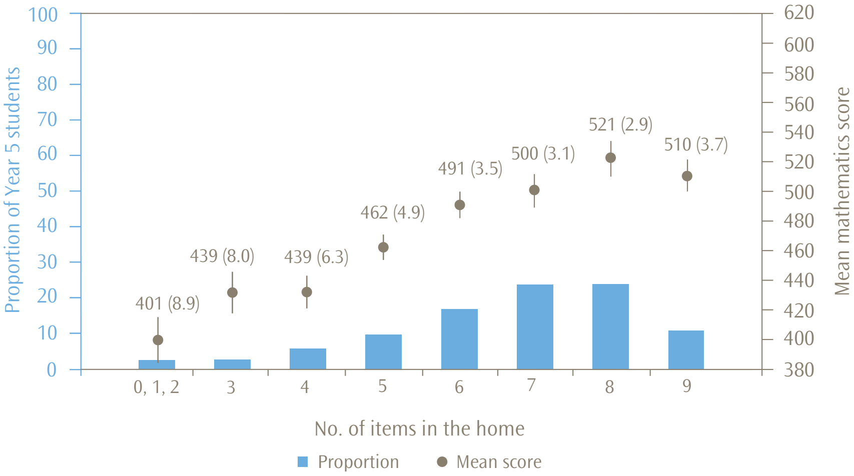
As Table 16 shows, the educational items were the most common items found in the homes of New Zealand Year 5 students. Approximately nine out of every ten students had a calculator in their home and a similar proportion reported having a computer in their home. Just over three-quarters of students reported an internet connection in their home and similarly, three-quarters reported that they had their own room. It was least common for students to have their own mobile phone with only 36 percent of students reporting this.

Table 16 **Proportion of students reporting item is in the home**

Item	Proportion of Year 5 students (%)
Calculator	92
Computer	91
Dictionary	89
Study desk/table for your own use	80
Internet connection	77
Your own room	77
Dishwasher	69
Musical instruments	62
Your own mobile phone	36

Eleven percent of students reported that all nine items could be found in their homes, one-quarter reported eight items, and a further quarter reported seven items. Just under 40 percent of students reported six or fewer of the items could be found in their homes, with less than one percent reporting one or none of the listed items. As shown in Figure 14, mathematics achievement generally increased as the number of items in the home increased.

Figure 14 Proportions and mean mathematics achievement of Year 5 students by number of items found in the home



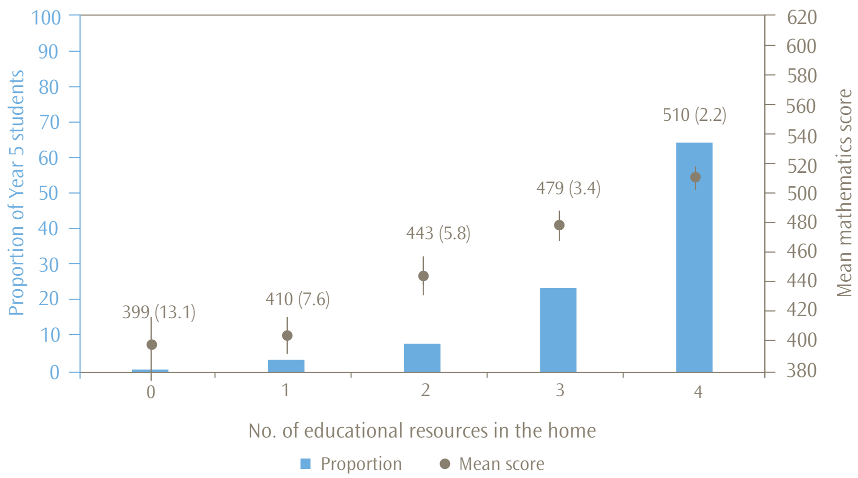
Note: *The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.*

Home educational resources

As mentioned earlier, the *calculator, computer (do not include PlayStation®, GameCube®, XBox®, or other TV/video game computers),* and *study desk/table for your use* were included in the list of items to ascertain the availability of educational resources at home. While students may not necessarily use these items for educational purposes, their presence could indicate the relative importance of education to the family, although this also may be reflective of the wealth of the home.

Nearly two-thirds of students (64%) reported that they had all four educational resources in their homes, while nearly one-quarter of students (24%) reported three of the four items in their homes. As shown in Figure 15, students with more educational resources in the home had higher mean mathematics achievement than those with fewer items.

Figure 15 Proportions and mean mathematics achievement of Year 5 students by number of educational resources found in the home



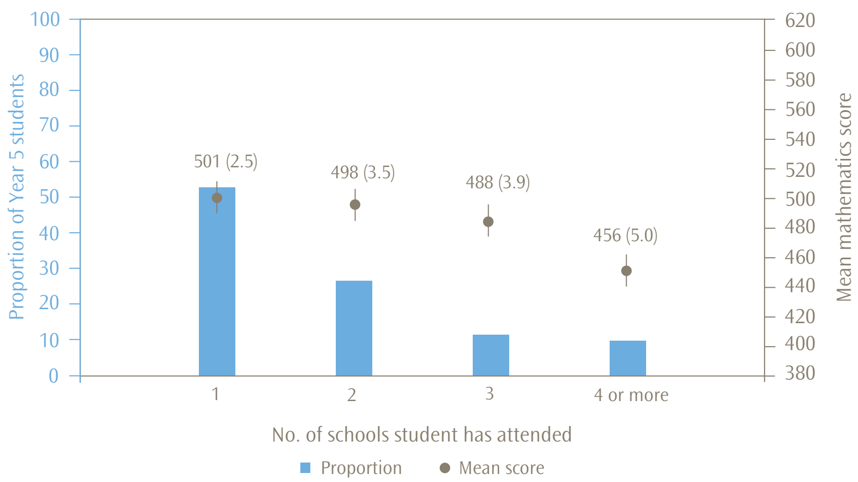
Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. The achievement value for the students with 0 educational resources should be treated with caution as indicated by the high standard error around the mean. Standard errors are presented in parentheses.

Number of schools attended by student

Another question included in the questionnaire that may be indicative of socio-economic status was the number of schools attended by a student. Many students in New Zealand change schools for a variety of reasons, but high mobility may be symptomatic of families moving regularly to find work, or students moving about among family members, or in care.

Just over half of all students reported that they had only attended one school, their current school, while one quarter of students had attended two schools. One in every 10 students reported they had attended four or more schools. As shown in Figure 16, students with high mobility had lower achievement than those with low mobility.

Figure 16 Proportions and mean mathematics achievement of Year 5 students by number of schools student has attended



Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.

Household size

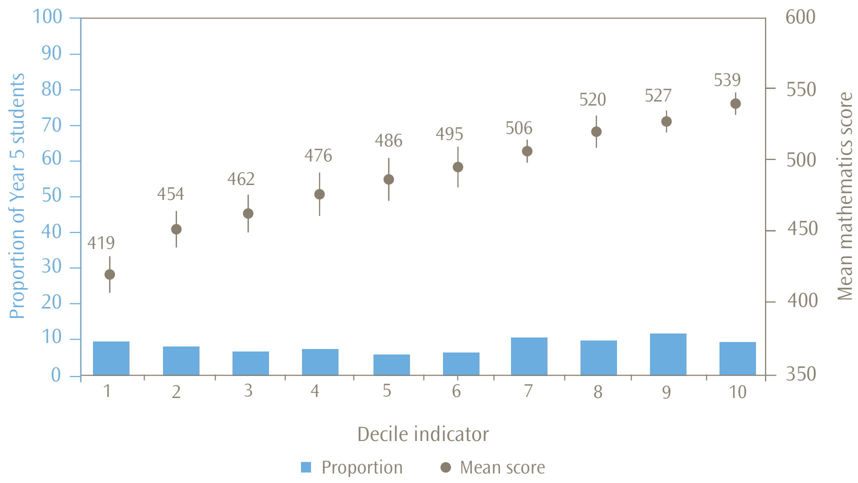
Another indicator of socio-economic status is the size of the household. While cultural or religious beliefs may determine household size, household crowding may also be indicative of poorer economic background. However, homes with one child living with one parent may also struggle financially. Mathematics achievement was examined with respect to number of people in the household. The highest achievement was found amongst students in households of size 3, 4, or 5 (501, 510, and 506 scale score points respectively) with the lowest achievement amongst students in households of 2, 7 or 8 (465, 466 and 437 respectively).

Decile

The Ministry of Education allocates resources such as Targeted Funding for Educational Achievement (TFEA) based on school decile indicator. A school's decile indicates the extent to which a school draws its students from low socio-economic communities. In general, decile 1 schools are the schools with the highest proportion of students from socio-economically disadvantaged communities, while decile 10 schools are the ten percent of schools with the lowest proportion of students from these communities.

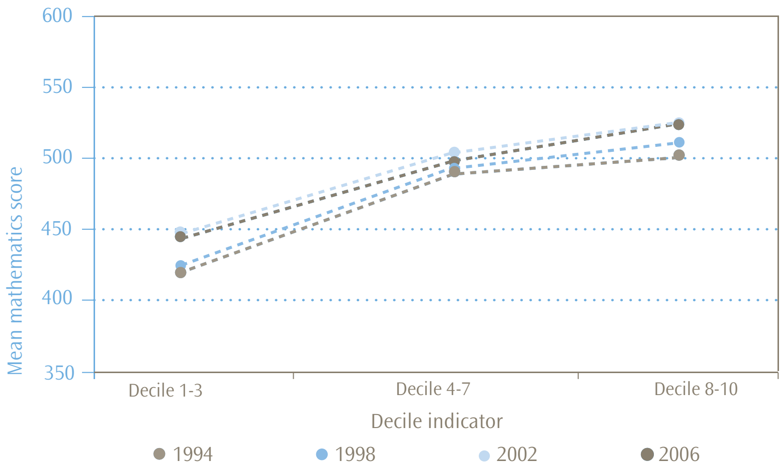
Analyses of mathematics achievement for students in schools in each decile band demonstrate that mathematics achievement is higher in higher decile schools and lower in lower decile schools as shown in Figure 17. However, the difference in achievement was not always significant when adjacent groups were examined. The largest difference between the mean mathematics score of adjacent groups occurs between students in deciles 1 and 2 (35 scale score points). It should be noted that this analysis does not demonstrate a causal link between being in a higher decile school and having higher achievement. Rather it is indicative of a trend demonstrating that students with lower levels of disadvantage in terms of family background and socio-economic background and living in wealthier communities have higher achievement.

Figure 17 Proportions and mean mathematics achievement of Year 5 students by decile indicator of school



Decile results were not reported in earlier TIMSS studies at this level. However, they have been analysed for this report by the broad groupings 1 to 3, 4 to 7, and 8 to 10 in order to ascertain whether the pattern of higher achievement, on average, among students in higher decile schools, was also evident in earlier cycles. As shown in Figure 18, students in higher decile schools have consistently demonstrated higher achievement than those in lower decile schools.

Figure 18 **Trends in mathematics achievement by decile indicator for 1994 to 2006**



Summary of mathematics achievement by socio-economic status and home educational resources

As the results in this section demonstrate, students from higher socio-economic backgrounds tend to have higher mean mathematics achievement than those from lower backgrounds as evidenced by the proxy measures *books in the home*, *items in the home*, *household size* and *mobility*. In addition, the decile of the school they attend, indicative of the level of economic disadvantage in the community in which they live, was positively related to mathematics achievement. That is students in higher decile schools had higher mathematics achievement, on average, than those in lower deciles.

Student activities outside of school

Previous cycles of TIMSS have shown that watching television and videos was the most popular leisure activity for Year 5 students (see Caygill, Sturrock, & Chamberlain 2007). Leisure activities are of interest in TIMSS as they can provide positive learning experiences as well as the negative implications of reducing time for doing school-related learning at home. While no judgements are made in this section of the value of leisure activities, and acknowledging that learning occurs within and outside of school, it is interesting to look at the changes over time and also to examine the relationships with achievement.

Table 17 presents the mean number of hours per school day that students reported spending on a variety of activities, along with the proportion of students who reported spending more than 2 hours on each activity. Note that it is possible that some of these activities were 'multi-tasked'. For example, students who spent an hour after school playing sport with friends selected *playing sports* and *playing or talking with friends*. Playing or talking with friends and playing sports were the two most popular activities for Year 5 students in 2006. Television watching was relegated to third most popular activity in 2006 in contrast with previous cycles of TIMSS.

Table 17 The proportion and mean amounts of time Year 5 students reported spending on leisure activities

Leisure activities	Mean number of hours per school day		Proportion spending more than 2 hours (%)
Watching television and videos	1.5	(0.03)	25
Playing computer games	1.0	(0.02)	15
Playing or talking with friends	1.7	(0.03)	31
Doing jobs at home	1.3	(0.03)	20
Playing sports	1.7	(0.03)	30
Reading a book for enjoyment	1.2	(0.02)	18
Using the internet	1.1	(0.02)	16

Note: Mean number of hours based on: No time=0; Less than 1 hour=0.5; 1-2 hours=1.5; More than 2 but less than 4 hours=3; 4 or more hours=4.5. Activities are not necessarily exclusive.
Standard errors are in parentheses.

The relationship between the number of hours spent in the individual activities and mathematics achievement was relatively consistent across the activities. Students who reported a small or moderate amount of time in an activity generally had higher achievement than those who reported no time or many hours on the activity.

Student attitudes

Students were asked how much they agree with eight statements about learning mathematics (listed in Table 18 – positive and negative statements were interwoven in the questionnaire but are reordered here for easier reading). They were given four response options: *agree a lot, agree a little, disagree a little, disagree a lot*.

Students were generally positive about mathematics with 79 percent agreeing that they enjoy learning mathematics and 68 percent disagreeing that mathematics was boring. A reasonably high proportion of the students agreed that they would like to do more mathematics in school (64%), but just under half the students agreed that mathematics was harder for them than for many of their classmates.

Table 18 Proportion of students who responded positively to statements about learning mathematics

Statements about learning mathematics	Proportion of students			
Positive statements	**Agreeing (%)**		**Disagreeing (%)**	
I usually do well in mathematics	86	(0.6)	14	(0.6)
I would like to do more mathematics in school	64	(1.0)	36	(1.0)
I enjoy learning mathematics	79	(0.9)	21	(0.9)
I learn things quickly in mathematics	74	(0.7)	26	(0.7)
I like mathematics	75	(0.8)	25	(0.8)
Negative statements	**Agreeing (%)**		**Disagreeing (%)**	
Mathematics is harder for me than for many of my classmates	47	(0.9)	53	(0.9)
I am just not good at mathematics	36	(0.8)	64	(0.8)
Mathematics is boring	32	(0.9)	68	(0.9)

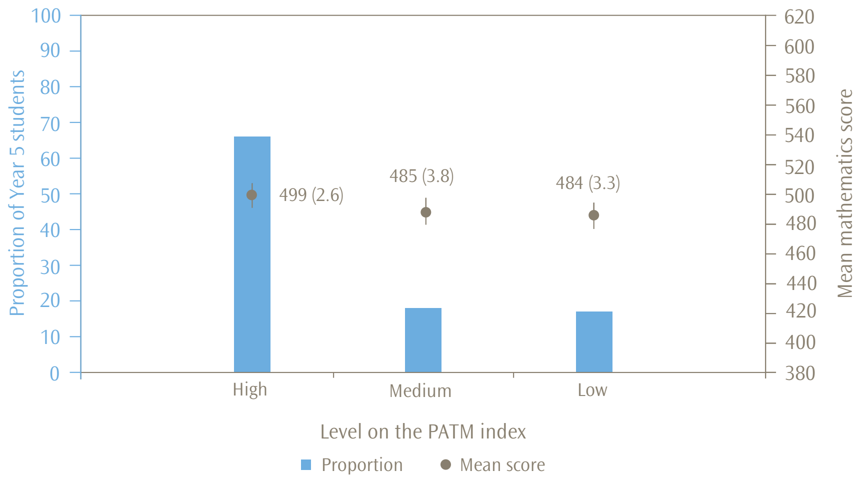
Note: The values for agree combine student responses to 'agree a lot' and 'agree a little'. Similarly the values for disagree combine little and a lot. Standard errors are presented in parentheses.

Generally, students with positive attitudes towards mathematics had higher achievement than students with negative attitudes. In order to examine the relationship with achievement the international researchers combined the data in two indices: the index of students' positive affect toward mathematics and the index of students' self-confidence in learning mathematics.

Index of students’ positive affect toward mathematics

The three statements: I enjoy learning mathematics; mathematics is boring; and I like mathematics; were combined to form the index of students’ positive affect toward mathematics (PATM).[15] Two-thirds of the students were at the high level of this index; that is, on average, they were positive about mathematics. Seventeen percent of students were at the low level of the index; that is, on average, they were negative about mathematics. There are now proportionally fewer students at the high level on the index and more at the medium and low levels than there were in 1994. As shown in Figure 19 students who were more positive about mathematics (at the high level of the PATM index) had higher mean mathematics achievement than those that were more negative. There was no difference between those at the medium and low levels of the index.

Figure 19 Proportion and mean mathematics achievement of students at each level of the positive affect toward mathematics (PATM) index



Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.

In relation to other countries, relatively low proportions of New Zealand students were at the high level of the PATM index. That is, few students reported positive attitudes towards mathematics in comparison to their international counterparts. However, this proportion was the same as the United States and Australia (66% each); England and Scotland had lower proportions of students who responded positively to these statements (62% and 59% respectively at the high level of the PATM index).

Index of students’ self-confidence in learning mathematics

The four statements: I usually do well in mathematics; mathematics is harder for me than for many of my classmates; I am just not good at mathematics; and I learn things quickly in mathematics; were combined to form the index of students’ self-confidence in learning mathematics (SCM).[16] Just over half (52%) of the students were at the high level of this index; that is, on average, they were positive about their own abilities in mathematics. Eleven percent of students were at the low level of the index; that is, on average, they were negative about their abilities in mathematics.

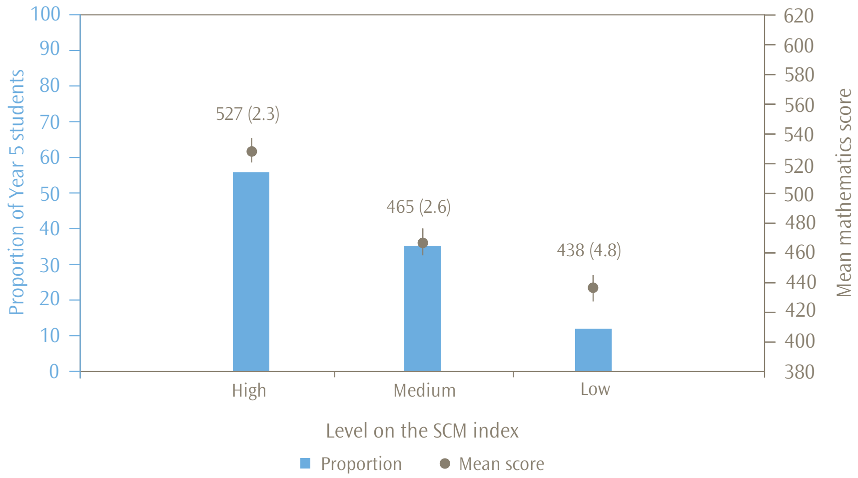
The proportions of the students at all levels of the index have changed significantly since 2002. More students are now positive about their abilities to learn mathematics (13 percentage point increase), but also more students are negative about their abilities to learn mathematics (7 percentage point increase). Fewer students are, therefore, at the medium level of the index.

15 An average was computed across a 4-point scale with 1 agree a lot, 2 agree a little, 3 disagree a little, 4 agree a lot. The statement ‘mathematics is boring’ was reversed so that students disagreeing a lot were given a value of 1.

16 An average was computed across a 4-point scale with 1 agree a lot, 2 agree a little, 3 disagree a little, 4 agree a lot. The statements ‘mathematics is harder for me than for many of my classmates’ and ‘I am just not good at mathematics’ was reversed so that students disagreeing a lot were given a value of 1.

As shown in Figure 20, students who were more positive about their abilities to learn mathematics (at the high level of the SCM index) had higher mean mathematics achievement than those that were more negative. Those students with the lowest self-confidence had the lowest mathematics achievement on average. Note that the difference in mean mathematics score between students that were high and those that were low on the SCM index (89 scale score points) is greater than those in the respective groups on the PATM index (15 scale score points). Thus the self-confidence of students had a stronger relationship with mathematics achievement than having a positive attitude towards mathematics.

Figure 20 Proportion and mean mathematics achievement of students at each level of the students' self-confidence in learning mathematics (SCM) index



Note: The bars on the graph represent the proportions of Year 5 students while the points represent mean scores. Lines extending from the points represent the 95% confidence interval, i.e. the range within which we are 95 percent confident that the true population value lies. Standard errors are presented in parentheses.

In relation to other countries, quite low proportions of New Zealand students were at the high level of the SCM index. However this proportion was similar to that for the Russian Federation (54%). In comparison, England, Australia, Scotland and the United States had a larger proportion of students who responded positively to these statements about their abilities in mathematics (64%, 64%, 67% and 67% respectively at the high level of the SCM index). High-performing countries including Singapore, Hong Kong and Japan had lower proportions at the low level of the SCM index (46%, 46% and 45% respectively).

Attitudes to mathematics by gender

Boys and girls demonstrated some similar attitudes to mathematics. Around two-thirds of girls and boys were very positive about mathematics and were at the high level of the PATM index (65% and 66% respectively). However, boys were more confident in their mathematics ability than girls, with a significantly higher proportion of boys at the high level of the SCS index (54% compared with 49% for girls).

Attitudes to mathematics by ethnicity

Some differences were evident among the ethnic groupings when attitudes to mathematics were considered. More Asian students (79%) reported positive attitudes to mathematics and were at the high level of the PATM index. The Other grouping had a similar proportion to Asian students with 78 percent. Significantly fewer Pākehā/European students (61%) reported positive attitudes to mathematics than any of the other ethnic groups (Pasifika 76%, and Māori 68%).

Similar to the PATM index, more Asian students were confident in their own mathematics abilities (65%) and were at the high level of the index. In contrast, fewer Pasifika (43%) and Māori (46%) students expressed high self-confidence compared to their Asian, Pākehā/European (53%), and Other (55%) counterparts.

Discussion of interactions

This report so far has presented results related to the mathematics achievement of New Zealand Year 5 students from TIMSS 2006/07, mainly in the form of descriptive statistics. The focus has been on background characteristics of the students in isolation from other characteristics, when in fact, many of these characteristics are interrelated. For example, one could speculate that differences in achievement between students of different ethnic groups are interrelated with differences in home language. Indeed, earlier in this report it was demonstrated by analysing these factors together, that the difference in mathematics achievement between those who spoke English at home regularly and those who did not was smaller when ethnic differences and immigration status were taken into account.

To investigate the possible interactions between characteristics of students and mathematics achievement outcomes further, some statistical modelling was undertaken. This involved putting all of the factors that have shown a relationship with achievement together in a statistical model using an analysis tool called MLWin.[17] This tool carries out multi-level modelling analysis of this type of data. A more detailed discussion of the modelling work will be presented in a separate working paper later in 2008 or early in 2009.

The value of looking at such models is that the relative importance of different background characteristics can be determined. Modelling also allows for the elimination of factors that are unimportant or measure the same underlying trait as others. The multi-level aspect of this type of statistical model takes account of the fact that student learning takes place within classes that are part of schools and that all of these can impact on achievement. Thus this type of model allows for consideration of differences in results between schools, classes, and students.

The first step of the analysis was to examine variations in mathematics achievement between schools, classes and students, without including any background characteristics. Around 20 percent of the variation in mathematics achievement was attributable to differences between schools, around 8 percent to differences between classes in the same school, and around 73 percent to differences between students. In other words, while there were some differences in mathematics achievement between the schools, and also the classes within schools, the majority of differences were between the students themselves.

In the second step of the analysis, the following background characteristics were considered in the model: gender, ethnicity, speaking English at home, born outside of New Zealand, student age, books in the home, possessions in the home, attitudes, out-of-school activities, school decile, mathematics instructional hours, and schools' participation in the Numeracy Project.

17 For further details see Goldstein (2003).

The model of influences on Year 5 students' mathematics achievement shows the following significant relationships with achievement when the other factors were taken account of:

- students with a greater number of books in the home had higher mathematics achievement than those who had fewer;
- older students had higher mathematics achievement than younger students;
- students born in New Zealand had higher mathematics achievement than those who were not;
- students with a greater number of educational resources in the home had higher mathematics achievement than those with fewer;
- students who spoke English frequently at home had higher mathematics achievement than those who did not;
- students with high self-confidence had higher mathematics achievement than those with lower self-confidence;
- students who felt safe at school had higher mathematics achievement than those who felt less safe;
- students who read books regularly had higher mathematics achievement than those who did not;
- boys had higher mathematics achievement than girls;
- Asian students had higher mathematics achievement than Pākehā/European students;
- Māori students had lower mathematics achievement than Pākehā/European students;
- Pasifika students had lower mathematics achievement than Pākehā/European students;
- students who were in a higher decile school (living in a community with less economic disadvantage) had higher mathematics achievement than those in lower decile schools;
- those students in a school that has participated in the Advanced Numeracy Project had higher achievement than those in schools that did not.

After completing the statistical model, variations between schools, classes, and students were re-examined to see if the model helped to explain the variations initially observed. The model explained just over sixty percent of the variation between schools, just under fifty percent of the variation between classes, and about thirty percent of the variation between students. In other words, the model explained much of the variation between schools, and between classes within schools, so that most of the variation that remained unexplained was between students.

Factors relating to the socio-economic status of the students such as decile, educational resources in the home, and books in the home all had a reasonably strong relationship with mathematics achievement in the model. While it is not easy to effect change in these, self-confidence in mathematics and reading books were two factors that had a strong relationship with achievement that might be influenced by teachers and parents.

This report has not examined all of the data collected in TIMSS. Further analyses will be undertaken including investigation of what is happening in mathematics in schools from information collected from teachers and principals. The model presented in this section is an initial investigation and should be read as such, but it gives valuable insight into the factors explaining differences in the mathematics achievement of Year 5 students in New Zealand.

Conclusion

This report has examined trends in New Zealand mathematics achievement at the Year 5 level from 1994 to 2006. It has looked at New Zealand Year 5 students' mathematics achievement in relation to other countries that participated in the study. An examination of the TIMSS assessment questions in relation to New Zealand's mathematics curriculum was presented along with analyses of achievement by sub-groupings (such as gender and ethnicity) and background factors. A statistical model that attempts to explain variations among students, classes, and schools using the background information discussed in this report was also described.

Achievement in mathematics

Overall, the mean mathematics achievement of New Zealand Year 5 students has risen between 1994 and 2006. In terms of the distribution of mathematics achievement across the range of scores, this is narrower in 2006 than in 1994. The positive aspect of this change is that fewer students are demonstrating very low achievement, while a similar proportion of New Zealand students are gaining very high scores. In international terms, New Zealand Year 5 mathematics achievement is significantly higher than 12 of the 36 TIMSS countries participating at the middle primary level, but significantly lower than 19 of the 36 countries.

Year 5 students continue to demonstrate relative strengths in aspects of mathematics. They tend to perform relatively better on *data display* questions compared to *number*. Students also perform relatively better on questions that involve *reasoning* compared to questions that assess knowledge.

Background characteristics

Both high and low performers were found among boys and girls, and in all ethnic groupings. On average, there was no difference in mathematics achievement between boys and girls. However, some differences were observed among the ethnic groupings. Asian students achieved higher in mathematics than Pākehā/European and the mean mathematics achievement of both these groups was higher, on average, than that of Māori and Pasifika students. Māori students had higher mean mathematics achievement than Pasifika students.

In terms of other background characteristics, mathematics achievement was higher, on average, among students who regularly spoke English at home. Similarly, students who were born in New Zealand had higher mathematics achievement on average than those who were not. Students from higher socio-economic backgrounds tended to have higher mean mathematics achievement than those from lower socio-economic backgrounds as evidenced by the proxy measures *books in the home*, *items in the home*, *household size* and *mobility*. In addition, the decile of the school they attended, indicative of the level of economic disadvantage in the community in which they live, was positively related to mathematics achievement. That is, students in higher decile schools had higher achievement, on average, than students in lower decile schools. Year 5 students who reported a small or moderate amount of time in out-of-school leisure activities generally had higher achievement than those who reported no time or many hours on the activity.

Attitudes to mathematics

New Zealand Year 5 students generally expressed positive attitudes towards mathematics. Those students who reported positive attitudes towards mathematics or were confident in their own mathematics abilities had higher achievement than those who were less positive or confident.

Boys and girls expressed similar attitudes to mathematics, in terms of enjoyment and motivation. However, boys were more confident in their mathematics ability than girls. More Asian students reported high self-confidence in mathematics compared with Pākehā/European, Māori, and Pasifika students and students in the Other ethnic grouping. More Asian students and students in the Other ethnic grouping reported positive attitudes towards mathematics compared with Pākehā/European, Māori, and Pasifika students.

Final comment

The Ministry of Education's current focus is on presence, engagement and achievement (Ministry of Education, 2007). The Ministry has in place a number of mechanisms for monitoring student performance in mathematics in primary and secondary schooling at the system level. Performance in mathematics reflects learning from within school, in family and whānau, and in the broader community. While overall there has been an increase in mean mathematics achievement between 1994 and 2006, variations in achievement among groupings require further attention.

This report has not examined all of the data collected in TIMSS. Further analyses will be undertaken including investigation of what is happening in mathematics in schools from information collected from teachers and principals. Further reports will become available during 2009.

References

Caygill, R., & Chamberlain, M. (2005). *Reading achievement in New Zealand in 1990 and 2001: results from the Trends in IEA's Reading Literacy Study*. Wellington: Ministry of Education.

Caygill, R. (2008). *Trends in Year 5 science achievement 1994 to 2006: results from the Trends in International Mathematics and Science Study (TIMSS)*. Wellington: Research Division, Ministry of Education.

Caygill, R., Sturrock, F., & Chamberlain, M. (2007). *Mathematics and science achievement in New Zealand: tracking the changes of Year 5 students in TIMSS 1994-2002*. Wellington: Ministry of Education.

Chamberlain, M. (2008). *PIRLS 2005/2006 in New Zealand: an overview of national findings from the second cycle of the Progress in International Reading Literacy Study*. Wellington: Ministry of Education.

Geske, A., Grinfelds, A., & Ozola, A. (2008). Latvia. In I.V.S. Mullis, M.O. Martin, J.F. Olson, D.R. Berger, D. Milne, & G.M. Stanco (Eds.), *TIMSS 2007 encyclopedia: a guide to mathematics and science education around the world* (pp. 347-354). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Goldstein, H. (2003). *Multilevel Statistical Models*. 3rd ed. London: Arnold.

Leung, F.K.S. &. Leung, C. C. Y. (2008). Hong Kong SAR. In I.V.S. Mullis, M.O. Martin, J.F. Olson, D.R. Berger, D. Milne, & G.M. Stanco (Eds.), *TIMSS 2007 encyclopedia: a guide to mathematics and science education around the world* (pp. 227-239). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Japelj Pavešić, B., & Svetlik, K. (2008). Slovenia. In I.V.S. Mullis, M.O. Martin, J.F. Olson, D.R. Berger, D. Milne, & G.M. Stanco (Eds.), *TIMSS 2007 encyclopedia: a guide to mathematics and science education around the world* (pp. 561-572). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Lam, L. (2002). *Mathematics education reform in Hong Kong*. Department of Mathematics, Hong Kong Institute of Education. Retrieved 30 October 2008 from: http://math.unipa.it/~grim/SiLam.PDF.

Martin, M.O., Mullis, I.V.S., & Foy, P. (2008). *TIMSS 2007 international science report: findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Ministry of Education (1992). *Mathematics in the New Zealand Curriculum*. Wellington: Learning Media Ltd.

Mullis, I.V.S., Martin, M.O., Ruddock, G.J., O'Sullivan, C.Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (2008). *TIMSS 2007 international mathematics report: findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). *TIMSS 2007 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Ruddock, G. (2008). England. In I.V.S. Mullis, M.O. Martin, J.F. Olson, D.R. Berger, D. Milne, & G.M. Stanco (Eds.), *TIMSS 2007 encyclopedia: a guide to mathematics and science education around the world* (pp. 181-191). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Tse, S.K. & Loh, E.K.Y. (2007). The impact of PIRLS in Hong Kong, SAR. In K. Schwippert (Ed.), *Progress in reading literacy: the impact of PIRLS 2001 in 13 countries*. Mūnster: Waxmann.

United Nations Development Programme (2008). *Human Development Report 2007/2008*. (pp. 229-232).

Definitions and technical notes

This section gives a brief overview of the technical details and definitions applicable to this report. For a comprehensive description of the technical details pertaining to TIMSS see the *TIMSS 2007 technical report* (Olson, Martin, & Mullis, (Eds.), 2008)

Benchmarks

In order to describe more fully what achievement on the mathematics scale means, the TIMSS international researchers have developed benchmarks. These benchmarks link student performance on the TIMSS mathematics scale to performance on mathematics questions and describe what students can typically do at set points on the mathematics achievement scale. The international mathematics benchmarks are four points on the mathematics scale, the advanced benchmark (625), the high benchmark (550), the intermediate benchmark (475), and the low benchmark (400). The performance of students reaching each benchmark is described in relation to the types of questions they answered correctly.

Exclusions

Each country was permitted to exclude some students for whom the assessment was not appropriate or was difficult to administer. Countries were required to keep the amount of excluded students as small as possible, with a guideline of 5 percent of the 'target' population as the maximum. Any countries that exceeded this value are indicated in the international exhibits. The target population in New Zealand was Year 5 students.

School-level exclusions in New Zealand consisted of very small schools (less than 4 Year 5 students), special education schools, Rudolf Steiner schools, the Correspondence School, and schools that provide more than 80% of their instruction in te reo Māori. Within-school exclusions consisted of special education classes, special needs students, students with insufficient instruction in English, and units within schools that provide more than 80% of their instruction in te reo Māori.

The New Zealand exclusion rate was one of the largest at 5.4 percent and equivalent to Hong Kong SAR and Lithuania. Exclusion rates for most of the other countries were usually kept below the 5 percent maximum, with only the United States and the benchmarking participants exceeding this level.[18]

Making models

The models in this report were formulated using two different methods. Regression analyses were used for the model at the student level that combined ethnicity, speaking English at home, and immigrant status. Custom-written programs described in the *TIMSS user guide for the international database* (to be published in early 2009) were used for this analysis. Multi-level modelling techniques were applied using the MLWin package for the analysis which examined school-, class-, and student-level variations in achievement. A range of background characteristics were included in the larger model initially and the model was then tested iteratively. At each iteration, any characteristics that were not statistically significant were removed until the model contained only variables with a significant influence on student achievement.

18 See Mullis, Martin, & Foy. (2008), Exhibit A.4 for this information.

Mean, medians, and averages

There are three measures of central tendency, but only the mean and the median are used in this report.

The mean of a set of scores is the sum of the scores divided by the number of scores, and is also sometimes referred to as 'the average', particularly in the international reports. Note that for TIMSS, as with other large-scale studies, the means for a country are adjusted slightly (in technical terms 'weighted') to reflect the total population of Year 5 rather than just the sample.

A median is the middle number when all numbers are put in order.

In earlier cycles of TIMSS, an international mean was reported. However as the number of countries participating changed, this mean shifted so that it was difficult to make comparisons across years. In TIMSS 2006/07 the TIMSS scale average is reported. This is the value to which the scores of each student are scaled (see later note on *Scale score points* for more details).

Minimum group size for reporting achievement data

In this report, student achievement data is not reported where the group size is less than 30 students or less than 10 schools. While group sizes of 30 to 50 students do have achievement reported in some cases, these are annotated and should be treated with caution as there is a lot of uncertainty in the measurement, as demonstrated by larger standard errors.

Percentile

The percentages of students performing below or above particular points on the scale can be used to describe the range of achievement. The lowest outer limit of achievement reported in ranges is the 5th percentile – the score at which only 5 percent of students achieved a lower score and 95 percent of students achieved a higher score. The highest outer limit is the 95th percentile – the score at which only 5 percent of students achieved a higher score and 95 percent of students a lower score. Therefore 90 percent of the Year 5 student scores lie between the 5th and 95th percentiles.

Sampling

Schools are sampled in TIMSS with a probability proportional to the number of Year 5 students. In order to improve the precision of sampling, the schools were ordered by decile, level of urbanisation, and size, so that the schools selected better represented the population of schools in New Zealand. Within each school, classes were sampled with equal probability and all Year 5 students within each class were selected.

Scale score points

The design of TIMSS allows for a large number of questions to be used in mathematics and science; each student answers only a portion of these questions. TIMSS employs techniques to enable population estimates of achievement to be produced for each country even though a sample of students responded to differing selections of questions. These techniques result in scaled scores that are on a scale with a mean of 500 and a standard deviation of 100.

Significance tests

In this report, all the comparisons that have been made are tested for statistical significance using the *t* statistic, with the probability of making an incorrect inference set at 5 percent. To compare the means of two groups of students, the formula to generate the test statistics computed in this report is:

(1) $$t = \frac{\bar{X}_1 - \bar{X}_2}{se_{diff}}$$

The calculation of se_{diff}, the standard error of the difference, varies depending on whether the groups were sampled independently or not. If the means for two groups that were sampled independently are being compared, for example, boys' achievement in 1994 and 2006, then the standard error of the difference is calculated as the square root of the sum of the squared standard errors of each mean:

(2) $$se_{diff} = \sqrt{se_1^2 + se_2^2}$$

For most of the comparisons, this formula was not applicable and so the se_{diff} is computed more accurately by combining variances using custom-written SAS programs. However as a rough estimate, the above formula will give a similar result.

Note that in all calculations, unrounded figures are used in these tests, which may account for some results appearing to be inconsistent.

Standard error

Because of the technical nature of TIMSS, the calculation of statistics such as means and proportions has some uncertainty due to

(i) generalising from the sample to the total Year 5 school population, and
(ii) inferring each student's proficiency from their performance on a subset of questions.

The standard errors provide a measure of this uncertainty. In general, we can be 95 percent confident that the true population value lies within an interval of 1.96 standard errors either side of the given statistic. This confidence interval is represented in graphs by the lines extending in either direction from the points.

Statistically significant

In order to determine whether a difference between two means is actual, it is usual to undertake tests of significance. These tests take into account the means and the error associated with them. If a result is reported as not being statistically significant then, although the means might be slightly different, we do not have sufficient evidence to infer that they are different. All tests of statistical significance referred to in this report are at the 95 percent confidence level.

Weighting

Due to the use of sampling, weights need to be applied when analysing the TIMSS data. Weighting ensures that any information presented more closely reflects the total population of Year 5 students rather than just the sample. The TIMSS weighting takes into account school, class, and student level information and the overall sampling weight is a product of the school, class, and student weights.

First published in December 2008 by:
Comparative Education Research Unit
Research Division
Ministry of Education
PO Box 1666
Wellington
New Zealand

Email: research.info@minedu.govt.nz
Fax: 64-4-463 8312
Phone: 64-4-463 8000

This report is available from the Education Counts website:
www.educationcounts.govt.nz

ISBN 978-0-478-13966-2
ISBN 978-0-478-13967-9 (web)
RMR-915