CHAPTER 3: Science Achievement

KEY POINTS

- There was a small, but non-significant, increase in the mean achievement for Year 5 students in science between 1994 and 1998.
- The science performance of boys between 1994 and1998 improved significantly, however there was no change in the achievement level of girls.
- Year 5 Maori and Asian ethnic groupings recorded the largest increases (non-significant) in science achievement over the same four-year period.
- In 1994, the mean science score of Maori girls was significantly higher than for their male counterparts. By 1998, the mean score for Maori boys had improved to be no different from that of Maori girls.
- A significant increase in science achievement between 1994 and 1998 was reported for those Year 5 students where English was spoken only rarely in the home.

This chapter focuses on the overall performance of Year 5 students (at the standard 3 level) in science in 1998, compared with the science achievement of Year 5 students in 1994[1]. Towards the end of this chapter a selection of items are presented along with the proportions of students answering correctly. As was the case with mathematics it is important to remember that there are no international results included in this chapter.

REPORTING STUDENT SCIENCE ACHIEVEMENT RESULTS

As described in Chapter 1, three approaches are used in this report to describe Year 5 students' science achievement:

- Item response theory (IRT) scale scores to describe overall achievement.
- Mean percent correct scores to examine student achievement on *trend* and *released* items.
- The percentage of students who answered each individual science item correctly.

Standard errors (se) have been calculated using the Jackknife Repeated Replication procedure. A 95 percent confidence interval applies. See Technical Notes TN.1 and TN.2 in Appendix E for more detail on the statistical methods used.

TRENDS IN SCIENCE ACHIEVEMENT

As noted in Chapters 1 and 2, two sets of information are used to examine trends in Year 5 students' achievement. Firstly, the overall 1998 science achievement results, in the form of IRT scale scores, are compared with the IRT scale scores for 1994[2]. Secondly, mean percent correct scores on the *trend*, and in some instances, *released* sets are presented either in the chapter or in Appendix C. The *released* items were made available after TIMSS-94 and can be found, for example, in the Assessment Resource Bank of the NZCER[3]. While there is no evidence of systematic use of these items by New Zealand schools, it is possible that some students had previously been exposed to release items prior to TIMSS-98 testing. This then, is the principal reason that analyses of *trend* items have been reported — *trend* results substantiate overall findings when they are consistent and place caveats around overall findings where they are inconsistent.

OVERALL SCIENCE ACHIEVEMENT

Table 3.1 presents the overall mean science scores for Year 5 students for TIMSS-94 and TIMSS-98. There was a small, albeit non-significant, increase in mean achievement over the four years.

---

1 *TIMSS-94 involved middle-primary students at two class levels – standards 2 and 3 (or international grades 3 and 4). Years 4 and 5 students are mostly in these two class levels. TIMSS-98 assessed students at a class level equivalent to standard 3 level and involved mostly Year 5 students. The students involved in TIMSS-98 will hereafter be referred to as 'Year 5' students.*

2 *This approach was necessary because the design of the study incorporated multiple-matrix-sampling; that is, each student received only a sample of the test questions.*

3 *New Zealand Council for Educational Research.*

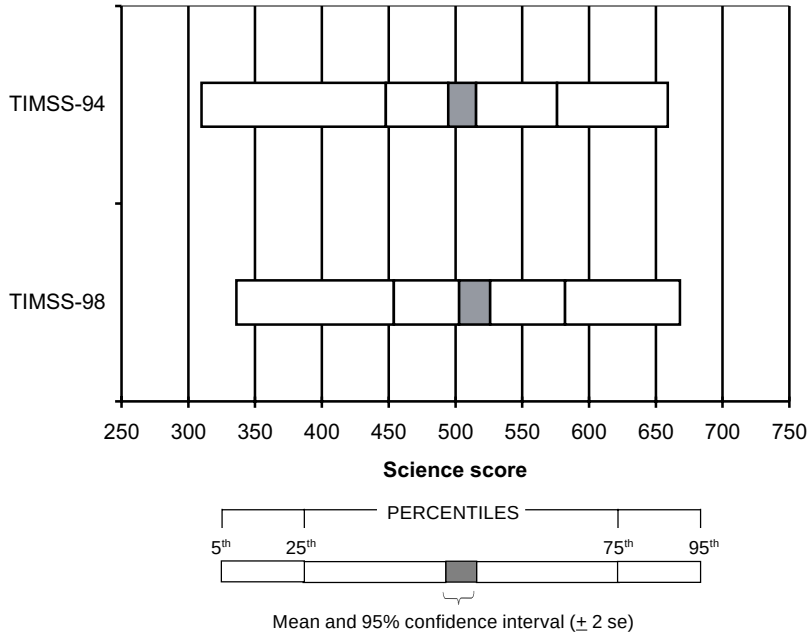
TABLE 3.1 YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1994 AND 1998

Year of assessment	Overall science mean (se)
1994	505 (5.3)
1998	514 (5.9)

Note: (se) Standard errors appear in parentheses.

Figure 3.1 presents information on the distribution of Year 5 students' science achievement scores for TIMSS-94 and TIMSS-98.

FIGURE 3.1 DISTRIBUTION OF YEAR 5 STUDENTS' SCIENCE SCORES FOR 1994 AND 1998



There were small, positive changes for all percentiles calculated for Year 5 students' science achievement from 1994 to 1998. For example, in 1994 the lowest achieving five percent of students scored below 310 but in 1998 this score had increased to 336. There was a much smaller increase in 95th percentile score, from 659 to 668. These results indicate a slightly smaller range of scores in 1998, as measured by the difference between the 5th and 95th percentiles. See Table C.1 in Appendix C for details of the calculated percentiles.

Trend item results

On an encouraging note, the small increase in Year 5 students' mean overall science achievement in 1998 was reflected in the increased mean score for the *trend* items, where the difference between 1994 and 1998 was of statistical significance (α=0.05). There was no change in the mean percent score on the *released* items over the same time period (see Table C.2 in Appendix C).

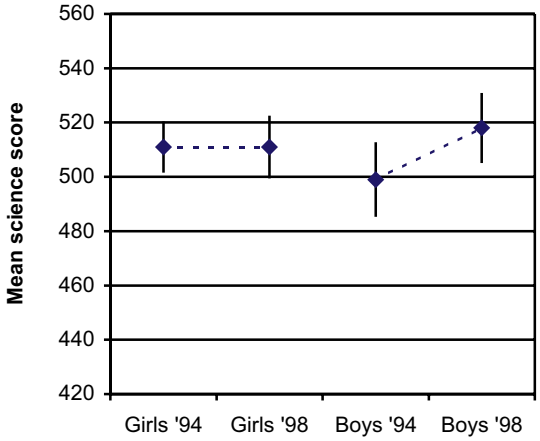
GENDER

Overall results

Figure 3.2 shows a breakdown of Year 5 students' mean science scores by gender for 1994 and 1998. Year 5 boys' mean science achievement increased by 19 scale score points between 1994 and 1998. This difference was of statistical significance (α=0.05). When this increase was examined in terms of effect size, it was found to be relatively small (d=0.17, see Technical Note TN.5 in Appendix E). The mean score for girls remained the same over the four-year period.

In 1994, the mean science score of Year 5 girls was 12 scale score points higher than that of Year 5 boys. In 1998 the converse was observed, whereby boys achieved a mean science score seven scale score points higher than their female counterparts. However, for both years, the difference between boys and girls' mean achievement was not of statistical significance. See Table C.3 in Appendix C for details.

FIGURE 3.2 YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1994 AND 1998, BY GENDER



Notes: See Chapter 4 for proportion of students by gender.

The data points are the mean scores. The vertical lines extending from the data points show the 95% confidence interval around the mean, ie ± 2 standard errors.

Trend item results

The increase in boys' overall mean science achievement was also reflected in the *trend* items. Boys' mean score on the *trend* item set increased by three percent from 70 percent in 1994 to 73 percent in 1998, with the difference of statistical significance (α=0.05). A similar improvement was evident for the boys on the *released* items (1994 = 54%; 1998 = 57%), but did not attain significance. There was little change in the girls' mean performance on either the *trend* or *released* items from 1994 to 1998. (See Table C.4 in Appendix C for details).

ETHNICITY

Overall results

Some interesting variations were evident in science achievement by ethnic grouping between 1994 and 1998. Table 3.2 presents the mean science scores for each of the main ethnic groupings for both 1994 and 1998. The largest increases in mean achievement from 1994 to 1998 were observed for students in the Maori and Asian groupings (21 and 24 scale score points respectively). However, due to the large standard errors[4] for both groupings, the differences between scores reported were not statistically significant.

[4] *Due mostly to the small sample sizes of these groups.*

TABLE 3.2: YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1994 AND 1998, BY ETHNIC GROUPING

Year of assessment	Mean science score (se)				
	Pakeha/ European	Maori	Pacific	Asian	Other
1994	534 (3.9)	457 (12.0)	441 (14.9)	493 (16.7)	521 (14.2)
1998	541 (4.8)	478 (8.0)	436 (13.8)	517 (10.0)	497 (23.0)

Notes:
[1] *See Chapter 4 for the proportions of students in each of the ethnic groupings.*
[2] *(s.e) Standard errors appear in parentheses.*

By way of contrast there was virtually no change in the mean scores for Pacific and Pakeha/European Year 5 students over the four-year period.

In 1994, the difference between the mean science scores for Pakeha/European and Maori was 77 scale score points; however in 1998, this difference was lower at 63 scale score points. The difference between Pakeha/European and Asian also reduced from 41 to 24 scale score points. In sharp contrast the difference between Pakeha/European and Pacific increased from 93 to 105 scale score points. How big was the shift in the relative performance of the groupings, in particular Maori and Pacific? As was the case for mathematics, effect sizes were calculated to illustrate the relative performance of students in each grouping over the four years. It is notable that the magnitude of the difference between the Pakeha/European and Maori groups reduced from a 'large' to a 'medium'-sized difference. Furthermore, the magnitude of the difference between Maori and Pacific students increased from 'small' to 'medium'. The magnitude of the difference between Pakeha/European and Pacific remained essentially unchanged — that is, a 'large' difference remains (see Tables C. 5a & C.5b in Appendix C, and Technical Note TN.5 in Appendix E for an explanation of effect size).

Trend item results

Performance on the *trend* items between 1994 and 1998 found that the improvement recorded by Asian (6 percentage points, α=0.05) and Maori (3 percentage points, not statistically significant) students were the only changes of note (see Table C.6 in Appendix C).

GENDER BY ETHNICITY

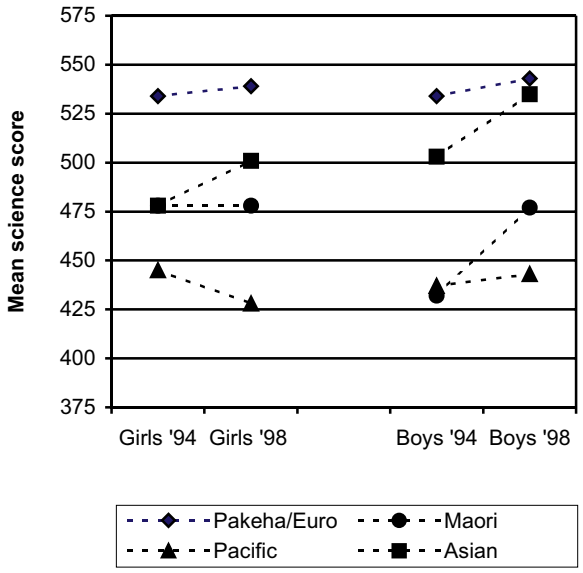
Overall results

Figure 3.4 shows a graphical representation of the science achievement of Year 5 girls and boys by ethnicity for 1994 and 1998. The most marked improvements over the four-year gap was observed for Maori boys — an increase of 45 scale score points (α=0.05), and to a lesser extent, Asian boys and Asian girls (32 and 23 scale score points respectively, not statistically significant). Pacific girls were the only sub-group whose mean science achievement decreased from 1994 to 1998 (17 scale score points, not statistically significant).

In 1994, one of the most significant gender differences existed in favour of Maori girls (46 scale score points, α=0.05) over their male counterparts. But the significant improvement in mean science achievement by Maori boys and the absence of change for Maori girls resulted in a virtually nil differential in mean achievement in 1998. Although not statistically significant, there were small increases in the

differential between girls' and boys' mean achievement over the intervening four years that favoured boys, in both the Pacific and Asian groupings. See Table C.7 in Appendix C for details.

FIGURE 3.3 YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1994 AND 1998, BY GENDER AND ETHNIC GROUPING



Trend item results

The patterns observed in overall scores were also observed in the set of science *trend* items (details are available in Table C.8 in Appendix C). That is, the largest increases in achievement over the four year period were observed for Maori boys (7 percentage points, α=0.05), Asian boys (7 percentage points, not statistically significant) and Asian girls (5 percentage points, not statistically significant). Pacific girls were the only group to record a decrease on *trend* item performance (3 percentage points, not statistically significant).

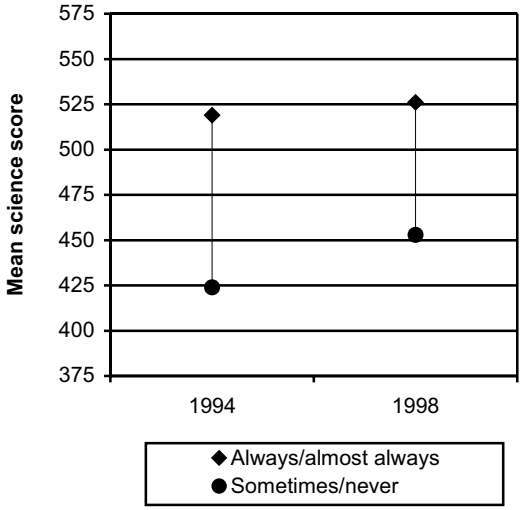
HOME LANGUAGE

Overall results

When the language spoken in the home is different from the language of the classroom, students face a disadvantage. Earlier IEA studies by Garden (1984), Lamb (1987), Wagemaker (1993), and Chamberlain G. (1997) have demonstrated that when the home language differs from that used for instruction by the teacher, the level of achievement is likely to be lower. Year 5 students were asked to indicate to what extent they spoke English at home and on the basis of their responses were divided into two groups: those who always or almost always spoke English at home, and those who only sometimes or never did. The mean science scores of these two groups for TIMSS-94 and TIMSS-98 are illustrated in Figure 3.4.

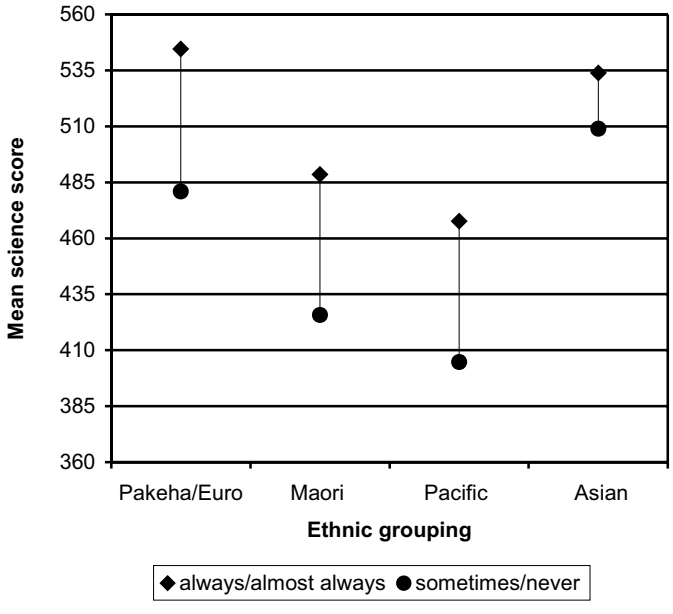
In 1998, the mean science score for students who spoke English at home *always or almost always* was 526 compared with a mean of 453 for those students who reported that they spoke English *sometimes / never*. The difference of 73 scale score points was statistically significant (α=0.05). By way of contrast, in 1994 the difference between the mean scores for students in these two home language groupings was 95 scale score points (α=0.05). Another illustration of the change of magnitude of this difference is that the effect size decreased from 0.96 in 1994 to 0.76 in 1998. Details of the figures are available in Table C.9 in Appendix C, see also Technical Note TN.5 in Appendix E.

FIGURE 3.4 YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1994 AND 1998, BY THE DEGREE THAT ENGLISH IS SPOKEN IN THE HOME



Of those Year 5 students who rarely spoke English at home, all ethnic groupings recorded some improvement in science achievement from 1994 to 1998, albeit small. However, only the increased science performance of Asian students was statistically significant (α=0.05). This pattern was similar to that described for mathematics earlier. The mean scores relating to the home language by ethnic grouping analyses can be found in Table C.10 in Appendix C. Figure 3.5 illustrates the mean science scores for those who frequently and rarely spoke English in the home, by ethnicity, for 1998.

FIGURE 3.5 YEAR 5 STUDENTS' MEAN SCIENCE SCORES FOR 1998, BY THE DEGREE THAT ENGLISH WAS SPOKEN IN THE HOME AND ETHNIC GROUPING



SCIENCE CONTENT REPORTING CATEGORIES

It is worth recalling from Chapter 1 that within each content domain, students were required to demonstrate a range of performances in order to answer the individual questions correctly. For example, approximately half of the *Earth Science* items required students to demonstrate *understanding* skills to answer correctly, while 36% of items in *Physical Science* were required to show *Theorising, analysing and solving problems* skills (for more detail see Adams & Gonzalez, 1996; Garden, 1997).

Overall results

This section provides a summary of the TIMSS-98 results focusing on the science content reporting categories of *Earth Science*, *Life Science*, and *Physical Science.* In some instances comparisons are made with information collected in 1994[5]. Due to the small number of items in some categories, there is no discussion on how students achieved on the *trend* items in any of the content reporting categories. Any change in performance should therefore be noted with caution. Note that as was the case for mathematics, the scores for each science reporting category were calculated separately from the overall science scale score and are therefore independent measures of achievement. These cannot be compared directly with the overall mean. However, in order to gauge areas of relative weakness or strength, it is possible to compare the mean for a given reporting category with the mean calculated for the sub-scales. The results for the science content reporting categories in 1994 and 1998 are presented in Figures 3.6a and 3.6b (see Table C.11 in Appendix C for further details).

FIGURE 3.6A YEAR 5 STUDENTS' MEAN SCORES FOR EACH SCIENCE REPORTING CATEGORY IN 1994

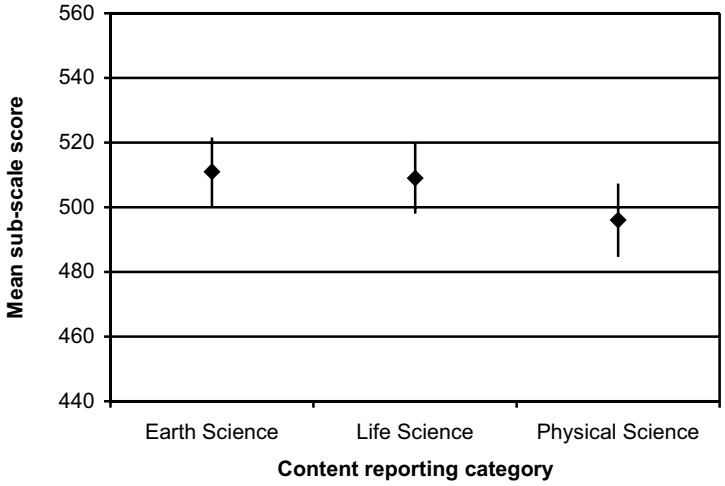
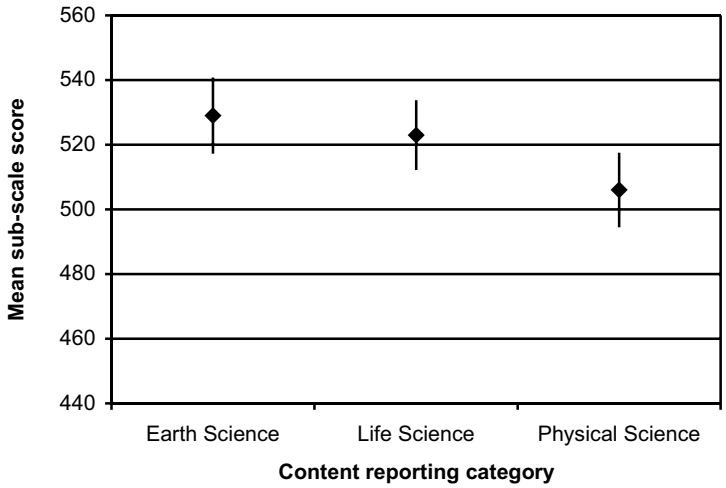


FIGURE 3.6B YEAR 5 STUDENTS' MEAN SCORES FOR EACH SCIENCE REPORTING CATEGORY IN 1998



The data points are the mean scores. The vertical lines extending from the data points show the 95% confidence interval around the mean, ie ± 2 standard errors.

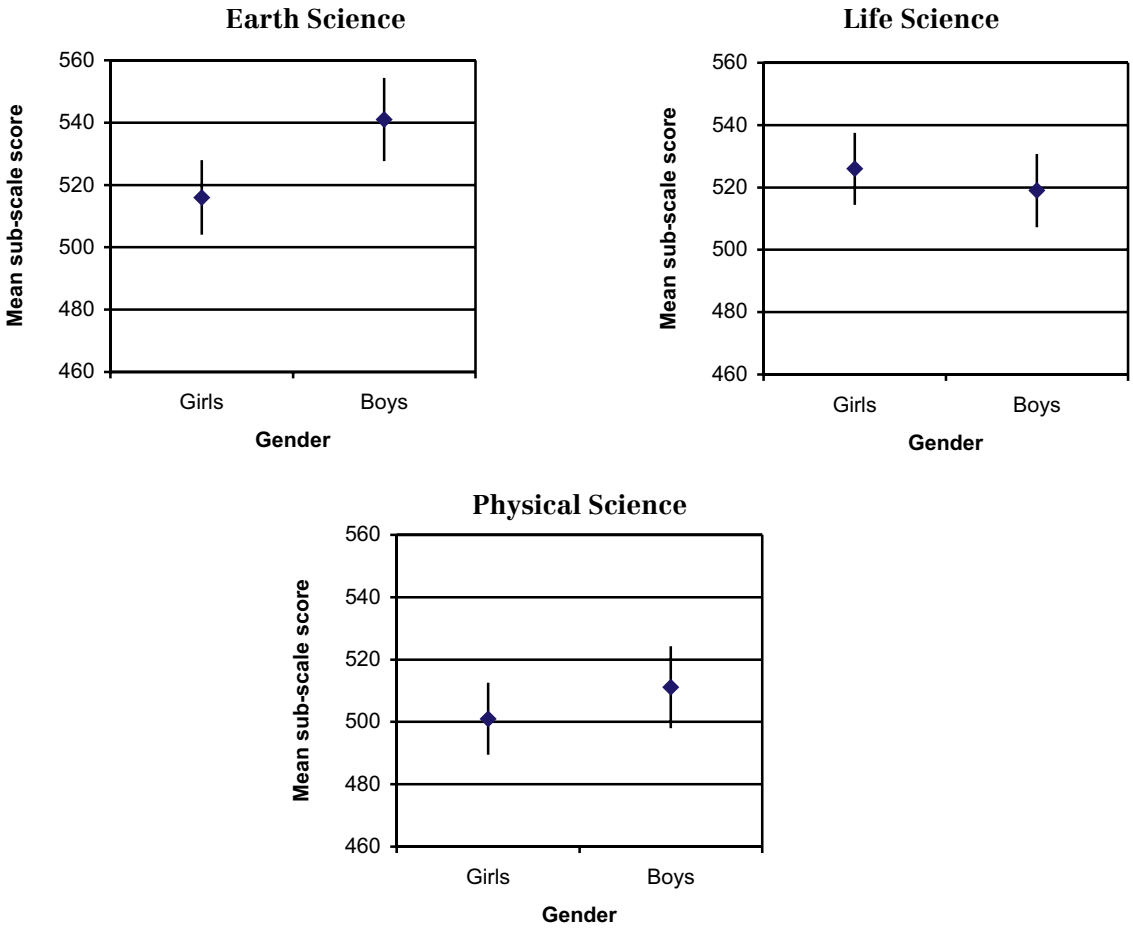
[5] *As noted in Chapter 1, the original TIMSS data collected across countries during 1994 and 1995 was re-scaled. To enable reliable sub-scale scores to be calculated for each of the reporting categories, items were subsequently regrouped into different categories from those first reported in TIMSS publications such as Martin et al (1997) and Garden (1997b). The discussion focuses on comparisons with the re-scaled 1994 achievement data and the scaled 1998.*

In 1994, there was less variation between the mean scores for each science content area relative to the mean for the sub-scale scores, although *Physical Science* was an area of weakness. However, in 1998 Year 5 students' performance was relatively stronger in *Earth Science* than in the other areas, while *Physical Science* was more clearly an area of relative weakness.

In general, Year 5 students made small, non-statistically significant gains across the four-year period in all three content reporting categories for science. The largest apparent growth was in *Earth Science* (18 scale score points). There were smaller improvements for *Life Science* (13) and *Physical Science* (10).

Figure 3.7 presents comparisons between boys' and girls' mean scores in each science content categories in 1998.

FIGURE 3.7 YEAR 5 STUDENTS' MEAN SCORES FOR EACH SCIENCE REPORTING CATEGORY IN 1998, BY GENDER



The data points are the mean scores. The vertical lines extending from the data points show the 95% confidence interval around the mean, ie ± 2 standard errors.

For *Earth Science* there was a statistically significant difference between boys and girls' mean scores of 25 scale score points (α=0.05) in favour of Year 5 boys in 1998. In *Life Science*, Year 5 girls' on average achieved scores that were not dissimilar to that of boys. By way of contrast, in 1994 girls achieved a significantly higher mean in *Life Science* than boys did (24 scale score points, α=0.05). The performance of Year 5 girls in 1998 in *Physical Science* was not dissimilar to that of boys. This was also

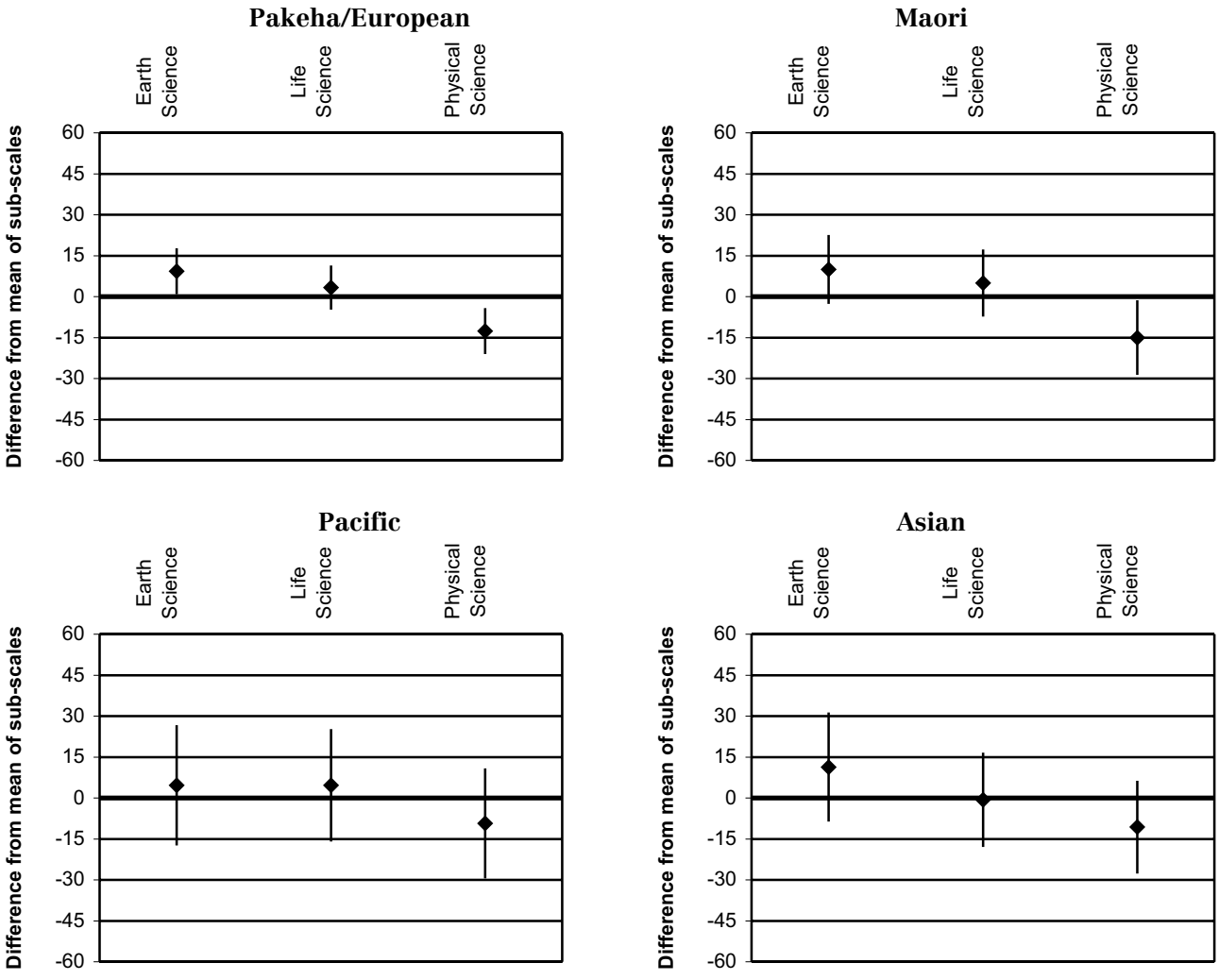
the case in 1994. Refer to Table C.12 in Appendix C for details of the mean scores for 1994 and 1998, by gender.

Comparing the performance across the science content areas by gender in 1998, both boys and girls were relatively weak in *Physical Science*. Boys' relative strength was in *Earth Science* whereas girls' relative strength was in *Life Science.*

From 1994 to 1998 there were no statistically significant changes in girls' mean scores in any of the science content reporting categories. However, boys' mean scores increased significantly in both *Earth Science* and *Life Science* (by 27 and 22 scale score points respectively, α=0.05).

Finally for the science content reporting categories, ethnicity is examined. Figure 3.8 compares the mean sub-scale scores for each ethnic grouping alongside the mean of these scores (see also Table C.13 in Appendix C). Note that the patterns of relative strength and weakness in content areas were similar across all ethnic groupings.

FIGURE 3.8 YEAR 5 STUDENTS' MEAN SCORES FOR EACH SCIENCE REPORTING CATEGORY IN 1998, BY ETHNIC GROUPING



The data points are the mean subscale scores. They are plotted by the difference from the mean of the three content area scores (represented by 0). The vertical lines extending from the data points show the 95% confidence interval around the score, ie ± 2 standard errors.

Between 1994 and 1998, there were consistent gains in mean science performance (23 to 28 scale score points) in all science content categories for Asian and Maori students, although these were not statistically significant. The performance of Pacific students was largely unchanged over the intervening four years. Pakeha/European students were the only ethnic grouping for which a statistically significant increase was observed in any science content area, this being 18 scale score points in *Earth Science* (α=0.05).

PERFORMANCE ON A SELECTION OF SCIENCE ITEMS

This final section of the science results focuses on how students achieved on a selection of individual *released* test questions. A (weighted) percentage is used to describe the proportion of students answering the questions correctly in both TIMSS-94 and TIMSS-98. Note that these items are reported as examples only. Systematic differences are more easily captured in the overall scale and content area sub-scale scores.

An example of one of the easiest science items for Year 5 students in 1994 and 1998.

Example 1

Content area: *Life Science (understanding complex information)*

R4. What is the MOST important reason for people to use sunscreen when they are outside in sunlight?

A. It protects the skin against dangerous rays from the sun.
B. It makes the skin more tanned.
C. It makes the skin smooth.
D. It makes the skin feel cooler.

1994 Year 5 % correct	1998 Year 5 % correct
81	87

An example of one of the difficult science items for Year 5 students in 1994 and 1998.

Example 2

Content area: *Physical Science (understanding complex information)*

R1. A watering can is almost filled with water as shown.

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

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

1994 Year 5 % correct	1998 Year 5 % correct
17	19

Examples of science items that exhibited an improvement for Year 5 students from 1994 to 1998

Example 3

Content area: *Earth Science (understanding simple information)*

O4. The Moon produces no light, and yet it shines at night. Why is this?

A. The Moon reflects the light from the Sun.
B. The Moon rotates at a very high speed.
C. The Moon is covered with a thin layer of ice.
D. The Moon has many craters.

1994 Year 5 % correct	1998 Year 5 % correct
64	73

Example 4

Content area: *Earth Science* (*understanding tools, routine procedures, and science process*)

N1. This table shows the temperature and precipitation (rain or snow) in four different towns on the same day.

	Town A	Town B	Town C	Town D
Lowest Temperature	13°C	-9°C	22°C	-12°C
Highest Temperature	25°C	-1°C	30°C	-4°C
Precipitation (rain or snow)	0 cm	5 cm	2.5 cm	0 cm

Where did it snow?

A. Town A
B. Town B.
C. Town C
D. Town D

1994 Year 5 % correct	1998 Year 5 % correct
39	46

Examples of science items demonstrating gender differences from 1994 to 1998

Example 5

Content area: *Life Science (understanding simple information)*

N6. A baby chick grows inside an egg for 21 days before it hatches. Where does the baby chick get its food before it hatches?

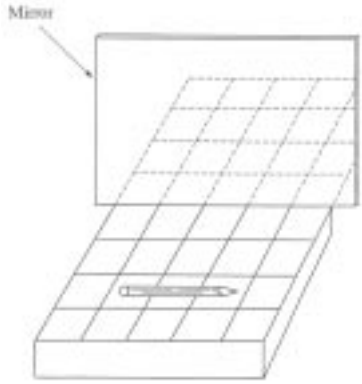
A. It is fed by the mother hen.
B. It doesn't need any food.
C. It makes its own food.
D. It uses food stored in the egg.
E. It eats the egg shell.

	1994 Year 5 % correct	1998 Year 5 % correct
Girls	51	32
Boys	48	43
All **students**	50	37

Example 6

Content area: *Physical Science (using tools, routine procedures, and science process)*

Q8. The picture shows a pencil that is lying on a shelf in front of a mirror. Draw a picture of the pencil as you would see it in the mirror. Use the patterns of lines on the shelf to help you.



	1994 Year 5 % correct	1998 Year 5 % correct
Girls	43	55
Boys	45	46
All **students**	44	50

Example 7

Content area: *Earth Science (theorising, analysing, and solving problems)*

Y1. The Sun is bigger than the Moon, but they appear to be about the same size when you look at them from the Earth. Why is this?

Year 5 students	1994 Year 5 % correct	1998 Year 5 % correct
Girls	31	34
Boys	36	49
All **students**	34	41