8 Performance of Subgroups

Although national monitoring has been designed primarily to present an overall national picture of student achievement, there is some provision for reporting on performance differences among subgroups of the sample. Eight demographic variables are available for creating subgroups, with students divided into subgroups on each variable, as detailed in Chapter 1 (p5).

Analyses of the relative performance of subgroups used an overall score for each task, created by adding together scores for appropriate components of the task.

SCHOOL VARIABLES

Five of the demographic variables related to the schools the students attended. For these five variables, statistical significance testing was used to explore differences in task performance among the subgroups. Where only two subgroups were compared (for *School Type*), differences in task performance between the two subgroups were checked for statistical significance using t-tests. Where three subgroups were compared, one-way analysis of variance was used to check for statistically significant differences among the three subgroups.

Because the number of students included in each analysis was quite large (approximately 450), the statistical tests were quite sensitive to small differences. To reduce the likelihood of attention being drawn to unimportant differences, the critical level for statistical significance was set at p = .01 (so that differences this large or larger among the subgroups would not be expected by chance in more than one percent of cases).

For the first four of the five school variables, statistically significant differences among the subgroups were found for less than 11 percent of the tasks at both year 4 and year 8. For the remaining variable, statistically significant differences were found on more than half of the tasks at both levels. In the detailed report below, all "differences" mentioned are statistically significant (to save space, the words "statistically significant" are omitted).

School Type

Results were compared for year 8 students attending full primary and intermediate schools. There were differences between these two subgroups on just two of 61 tasks.

Students from intermediate schools scored higher on *Soak It Up* (p44), but lower on *Link Task 18* (p45). There were also differences on two questions of the *Year 8 Science Survey* (p57), with students from full primary schools reporting greater experience of field trips or outside work, and more experience of research/projects.

School Size

Results were compared from students in large, medium-sized, and small schools (exact definitions were given in Chapter 1). For year 4 students, there was a difference among the subgroups on just one of 57 tasks: students from small schools scored highest on *Link Task 1* (p24). There were no differences on questions of the *Year 4 Science Survey* (p56).

For year 8 students, there were differences on two of the 61 tasks. Students from medium-sized schools scored highest (and those from small schools lowest) on *Link Task 22* (p45), while students from small schools scored highest on *Link Task 24* (p54). There were also differences on six questions of the *Year 8 Science Survey* (p57). Students from medium-sized schools were least keen to do more science at school, least positive about continuing to learn about science as they grew up, but felt that they most often did "really good things in science" at school. Students from small schools reported the most experience of field trips or outside work, the most experience of research/projects, and the least experience of experiments with science equipment (students from medium-sized schools were highest on that).

Community Size

Results were compared for students living in communities containing over 100,000 people (main centres), communities containing 10,000 to 100,000 people (provincial cities), and rural areas or towns containing less than 10,000 people (rural areas).

For year 4 students, there was a difference on just one of 57 tasks: students from main centres scored lowest on *Link Task 22* (p45). There were no differences on questions of the *Year 4 Science Survey* (p56).

For year 8 students, there were differences among the three subgroups on six of 61 tasks. Students from the main centres scored lowest on *Link Task 5* (p24), *Link Task 15* (p35), and *Link Task 24* (p54), but highest on *Shining Light* (p29). Students from provincial cities scored highest on *Link Task 5* (p24) and *Link Task 18* (p49). Students from rural areas scored lowest on *Shining Light* (p29) and *Link Task 25* (p54).

Zone

Results achieved by students from Auckland, the rest of the North Island, and the South Island were compared.

For year 4 students, there were differences among the three subgroups on six of 57 tasks. Students from Auckland scored highest on *Link Task 1* (p24) and *Scooter* (p33), but lowest on *Possums and Rabbits* (p16), *Rusty Tools* (p42), *Sand* (p48), and *Link Task 18* (p45). Students from the rest of the North Island scored lowest on *Link Task 1* (p24) and *Scooter* (p33). Students from the South Island scored highest on *Possums and Rabbits* (p16) and *Rusty Tools* (p42). There were no differences on questions of the *Year 4 Science Survey* (p56).

For year 8 students, there were differences among the three subgroups on three of the 61 tasks. Students from the South Island scored highest on *Link Task 15* (p35), *Link Task 22* (p45), and *Greenhouse Problem* (p51), with students from the North Island excluding Auckland lowest on the last of these. There were also differences on four questions of the *Year 8 Science Survey* (p57). Students from the South Island were least positive about studying science at school, doing more science at school, and doing science in their own time. Students from Auckland thought they learned the most about science at school.

Socio-Economic Index

Schools are categorised by the Ministry of Education based on census data for the census mesh blocks where children attending the schools live. The SES index takes into account household income levels, categories of employment, and the ethnic mix in the census mesh blocks. The SES index uses ten subdivisions, each containing ten percent of schools (deciles 1 to 10). For our purposes, the bottom three deciles (1-3) formed the low SES group, the middle four deciles (4-7) formed the medium SES group, and the top three deciles (8-10) formed the high SES group. Results were compared for students attending schools in each of these three SES groups.

For year 4 students, there were differences among the three subgroups on 37 of 57 tasks. Because of the large number of tasks involved, they will not be listed here. Students in high decile schools performed better than students in low decile schools on all 37 tasks, with students in medium decile schools somewhere between. There were no differences on questions of the *Year 4 Science Survey* (p56).

For year 8 students, there were differences among the three subgroups on 40 of 61 tasks. Because of the large number of tasks involved, they will not be listed here. Students in high decile schools performed better than students in low decile schools on all 40 tasks, with students in medium decile schools generally closer to their students in high decile schools. There were no differences on questions of the *Year 8 Science Survey* (p57).

STUDENT VARIABLES

Three demographic variables related to the students themselves:

Gender: boys and girls

Ethnicity: Māori, Pasifika, and Pakeha (this term was used for all other students)

Language used predominantly at home: English and other.

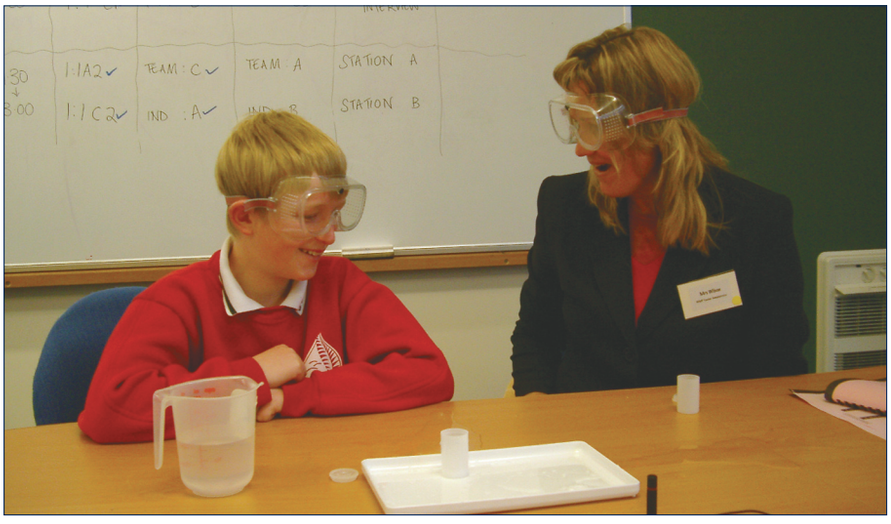
During the previous cycle of the Project (1999-2002), special supplementary samples of students from schools with at least 15 percent Pasifika students enrolled were included. These allowed the results of Pasifika students to be compared with those of Māori and Pakeha students attending these schools. By 2002, with Pasifika enrolments having increased nationally, it was decided that from 2003 onwards a better approach would be to compare the results of Pasifika students in the main NEMP samples with the corresponding results for Māori and Pakeha students. This gives a nationally representative picture, with the results more stable because the numbers of Māori and Pakeha students in the main samples are much larger than their numbers previously in the special samples.

The analyses reported here compare the performances of boys and girls, Paheha and Māori students, Pakeha and Pasifika students, and students from predominantly English speaking and non-English speaking homes.

For each of these three comparisons, differences in task performance between the two subgroups are described using "effect sizes" and statistical significance.

For each task and each year level, the analyses began with a t-test comparing the performance of the two selected subgroups and checking for statistical significance of the differences. Then the mean score obtained by students in one subgroup was subtracted from the mean score obtained by students in the other subgroup, and the difference in means was divided by the pooled standard deviation of the scores obtained by the two groups of students. This computed effect size describes the magnitude of the difference between the two subgroups in a way that indicates the strength of the difference and is not affected by the sample size. An effect size of +.30, for instance, indicates that students in the first subgroup scored, on average, three tenths of a standard deviation higher than students in the second subgroup.

For each pair of subgroups at each year level, the effect sizes of all available tasks were averaged to produce a mean effect size for the curriculum area and year level, giving an overall indication of the typical performance difference between the two subgroups.



Gender

Results achieved by male and female students were compared using the effect size procedures. Positive effect sizes indicate that boys did better on those tasks.

For year 4 students, the mean effect size across the 50 tasks was +.08 (boys averaged 0.08 standard deviations higher than girls). This is a small difference. There were statistically significant differences on seven of the 50 tasks. Boys performed better on all seven tasks: *Inside Outside Skeletons* (p19), *Shining Light* (p29), *Experimenting with Air and Water* (p38), *Rusty Tools* (p42), *Planets* (p52), *Landforms* (p53), and *Link Task 24* (p54). There were also differences on two questions of the

Year 4 Science Survey (p56): Boys reported doing more good things in science in their own time and were more positive about continuing to learn science as they grew up.

For year 8 students, the mean effect size across the 53 tasks was +.09 (boys averaged 0.09 standard deviations higher than girls). This is a small difference. There were statistically significant differences on 17 of the 53 tasks. Boys performed better on 14 of these tasks (1 living world, 6 physical world, 3 material world and 4 planet earth and beyond). Girls performed better on three tasks: *Plants* (p13), *Link Task 8* (p24), and *Link Task 22* (p45). There was also a difference on one question of the Y*ear 8 Science Survey* (p57): Boys were keener to do more science at school.

Ethnicity

Results achieved by Māori, Pasifika and Pakeha (all other) students were compared using the effect size procedures. First, the results for Pakeha students were compared to those for Māori students. Second, the results for Pakeha students were compared to those for Pasifika students. Positive effect sizes indicate that Pakeha students did better than the Māori or Pasifika students.

Pakeha-Māori Comparisons

For year 4 students, the mean effect size across the 50 tasks was +.30 (Pakeha students averaged 0.30 standard deviations higher than Māori students). This is a moderate difference. There were statistically significant differences on 20 of the 50 tasks, with Pakeha students performing better on all 20 tasks (10 living world, 2 physical world, 5 material world, and 3 planet earth and beyond). There were no differences on questions of the Year 4 Science Survey.

For year 8 students, the mean effect size across the 53 tasks was +.37 (Pakeha students averaged 0.30 standard deviations higher than Māori students). This is a moderate difference. There were statistically significant differences on 34 of the 53 tasks: Pakeha students performed better on these 34 tasks (12 living world, 6 physical world, 6 material world, and 10 planet earth and beyond). There was also a difference on one question of the *Year 8 Science Survey* (p57): Māori students reported greater involvement in research and project work in science at school.

Pakeha-Pasifika Comparisons

Readers should note that only 30 to 50 Pasifika students were included in the analysis for each task. This is lower than normally preferred for NEMP subgroup analyses, but has been judged adequate for giving a useful indication, through the overall pattern of results, of the Pasifika students' performance.

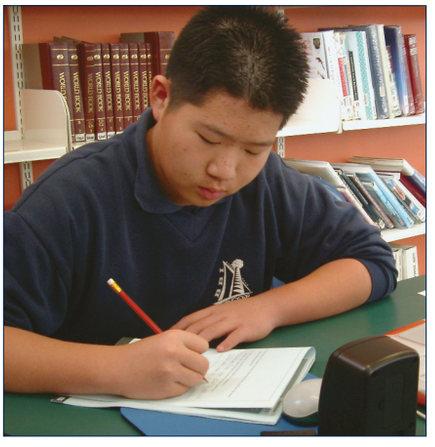
For year 4 students, the mean effect size across the 50 tasks was +.57 (Pakeha students averaged 0.57 standard deviations higher than Pasifika students). This is a large difference. There were statistically significant differences on 32 of the 50 tasks: Pakeha students performed better on all 32 tasks (12 living world, 4 physical world, 7 material world, and 9 planet earth and beyond). There were no differences on questions of the *Year 4 Science Survey* (p56).

For year 8 students, the mean effect size across the 53 tasks was +.62 (Pakeha students averaged 0.62 standard deviations higher than Pasifika students). This is a large difference. There were statistically significant differences on 37 of the 53 tasks: Pakeha students performed better on all 37 tasks (12 living world, 7 physical world, 9 material world, and 9 planet earth and beyond). There were no differences on questions of the Year 8 Science Survey.

Home Language

Results achieved by students who reported that English was the predominant language spoken at home were compared, using the effect size procedures, with the results of students who reported predominant use of another language at home (most commonly an Asian or Pasifika language. Positive effect sizes indicate that students for whom English was the predominant language at home performed better on those tasks.

For year 4 students, the mean effect size across the 50 tasks was +.37 (students for whom English was the predominant language at home averaged 0.37 standard deviations higher than the other students). This is a moderate difference. There were statistically significant differences on 26 of the 50 tasks: students for whom English was the predominant language spoken at home performed better on these 26 tasks (9 living world, 4 physical world, 6 material world, and 7 planet earth and beyond). There was also a difference on one questions of the *Year 4 Science Survey* (p56): students who predominant language at home was not English were more keen to do reading, viewing, writing, or listening activities related to science in their own time.



For year 8 students, the mean effect size across the 53 tasks was +.31 (students for whom English was the predominant language at home averaged 0.31 standard deviations higher than the other students). This is a moderate difference. There were statistically significant differences on 18 of the 53 tasks: students for whom English was the predominant language spoken at home performed better on these 18 tasks (6 living world, 0 physical world, 5 material world, and 7 planet earth and beyond). There was also a difference on one question of the *Year 8 Science Survey* (p57): students who predominant language at home was not English reported doing more good things in science in their own time.