

# 6 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 17 percent of the tasks at both year 4 and year 8. For the remaining variable, statistically significant differences were found on one third or more 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 no differences between these two subgroups on any of the 18 tasks. There were, however, differences on two questions of the *Year 8 Art Survey* (p47), with students from intermediate schools reporting greater experience of drawing and working with clay in art at school.

### 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 18 tasks: students from large schools scored highest on the monotype printing task *Dog Walk* (p25). There were no differences on questions of the *Year 4 Art Survey* (p46).

For year 8 students, there were no differences on any of the 18 tasks. There was a difference on one question of the *Year 8 Art Survey* (p47), with students from large schools highest and students from small schools lowest in reported experience of working with clay at school.

### 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 communities containing less than 10,000 people (rural areas).

For year 4 students, there were differences on two of the 18 tasks. Students from provincial towns scored lowest and students from Auckland highest on *Link Task 2* (p30), but the opposite was the case on *Link Task 9* (p44). There was a difference on one question of the *Year 4 Art Survey* (p46), with students from rural areas



reporting greater experience at school of looking at art and talking about it.

For year 8 students, there were differences among the three subgroups on two of the 18 tasks. Students from the main centres scored lowest on *Link Task 6* (p43), while students from rural areas scored lowest on *Link Task 9* (p44). There were no differences on questions of the *Year 8 Art Survey* (p47).

### 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 three of the 18 tasks. Students from Auckland scored highest and students from the South Island lowest on *Dog Walk* (p25) and *Link Task 2* (p30), but students from Auckland scored lowest on *Link Task 9* (p44). There were also differences on two questions of the *Year 4 Art Survey* (p46): students from the South Island were least positive about how often their class did really good things in art and how often they learned new things in art at school.

For year 8 students, there were differences among the three subgroups on two of the 18 tasks. Students from Auckland scored highest and students from the South Island lowest on *Dog Walk* (p25), but this advantage was reversed on *Link Task 6* (p43). There was also a difference on one question of the *Year 8 Art Survey* (p47). Somewhat ironically, given their poorer performance on *Dog Walk* (p25), students from the South Island reported the most experience of doing printmaking 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 10 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 six of the 18 tasks: *Link Task 3* (p31), *Warriors and Soldiers* (p33), *Art Objects* (p34), *Meeting House* (p36), *Wearable Arts* (p40), and *Link Task 6* (p43). Only one of these was an art making task. Students in high decile schools performed better than students in low decile schools on all six tasks, with students in medium decile schools generally closer to the students in low decile schools. There were also differences on three questions of the *Year 4 Art Survey* (p46), with students from low decile schools reporting more school experience of collage, carving and working with fabrics or weaving.

For year 8 students, there were differences among the three subgroups on nine of the 18 tasks: *Underwater Garden* (p14), *Link Task 1* (p30), *Warriors and Soldiers* (p33), *Art Objects* (p34), *Meeting House* (p36), *Portrait Pairs* (p38), *George Street* (p41), and *Link Task 8* (p44). The first two involved painting and drawing, with the other seven involving responding to art. Students in high decile schools performed better than students in low decile schools on all nine tasks, with students in medium decile schools generally closer to the students in high decile schools. There were also differences on six questions of the *Year 8 Art Survey* (p47), with students from high decile schools reporting learning least about art at school and spending the least time at school on drawing, working with fabric/weaving, and group activities in art. They also felt they learned the least new things in art, and were less inclined to believe that they would make good artists when they grew up.

## 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 per cent 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 compare the performances of boys and girls, Pakeha 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  $+0.30$ , 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. Because there was often a different pattern for the art making and responding to art tasks, mean effect sizes were also computed and reported for these two types of task.

### 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 17 tasks was  $-0.01$  (girls averaged 0.01 standard deviations higher than boys). This difference is negligible. On average, boys performed slightly better than girls on the responding to art tasks (mean effect size  $+0.05$ ), but girls performed a little better than boys

on the art making tasks (mean effect size  $-0.13$ ). There were statistically significant differences on two of the 17 tasks: girls performed better on *Underwater Garden* (p14) but boys performed better on *Wearable Arts* (p40). There were also differences on three questions of the *Year 4 Art Survey* (p44): Girls were more positive about doing art at school, how good their parents thought they were at art, and continuing to learn art as they grew up.

For year 8 students, the mean effect size across the 17 tasks was  $-0.09$  (girls averaged 0.09 standard deviations higher than boys). This is a small difference, but there was a slightly larger difference on the art making tasks (mean effect size  $-0.19$ ). There were statistically significant differences on three of the 17 tasks, with girls performing better on all three tasks: *Underwater Garden* (p14), *Bird Battle* (p20), and *Art Objects* (p34).



There were also differences on four questions of the *Year 8 Art Survey* (p47): Girls were more positive about doing art at school, about doing more art at school, and about doing art in their own time, but reported less experience of printmaking in art 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 17 tasks was +.31 (Pakeha students averaged 0.31 standard deviations higher than Māori students). This is a moderate difference. The difference was larger for responding to art tasks (+.38) than for art making tasks (+.17). There were statistically significant differences on seven of the 17 tasks (all of which were responding to art tasks): Pakeha students performed better on all seven tasks. There were differences on eight questions of the *Year 4 Art Survey* (p46): Māori students thought they learned more about art at school and had more opportunity at school to do drawing, carving, model making and work with fabrics or weaving. They also thought they had more opportunities at school to look at and talk about art and to learn new things about art, and a higher proportion thought they would make good artists when they grew up.



For year 8 students, the mean effect size across the 17 tasks was +.27 (Pakeha students averaged 0.27 standard deviations higher than Māori students). This is a moderate difference. The difference was larger for responding to art tasks (+.33) than for art making task (+.17). There were statistically significant differences on seven of the 17 tasks (*Dog Walk* (p25) and six responding to art tasks): Pakeha students performed better on all seven tasks. There were also differences on two questions of the *Year 8 Art Survey* (p47): Māori students reported more work with computer graphics at school and more often learning new things in art 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 17 tasks was +.37 (Pakeha students averaged 0.37 standard deviations higher than Pasifika students). This is a moderate difference. The difference was much larger for responding to art tasks (+.53) than for art making tasks (+.09). There were statistically significant differences on eight of the 17 tasks (all of which were responding to art tasks): Pakeha students performed better on all eight tasks. There were differences on 10 questions of the *Year 4 Art Survey* (p46): Pasifika students thought they learned more about art at school, more often did really good things in art at school, and had more opportunity at school to do drawing, printmaking, collage, carving, model making, and work with fabrics or weaving. They also thought they had more opportunities at school to learn new things about art, and a higher proportion wanted to keep learning about art when they grew up.

For year 8 students, the mean effect size across the 17 tasks was +.42 (Pakeha students averaged 0.42 standard deviations higher than Pasifika students). This is a moderate to large difference. The difference was substantially larger for responding to art tasks (+.53) than for art making tasks (+.21). There were statistically significant differences on eight of the 17 tasks (*Underwater Garden* (p14), *Link Task 1* (p30), and six responding to art tasks): Pakeha students performed better on all eight tasks. There were differences on five questions of the *Year 8 Art Survey* (p47): Pasifika students thought they learned more about art at school, more often did really good things in art at school, had more opportunity at school to do drawing and computer graphics, and learned more new things in art at school.

### 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 17 tasks was +.26 (students for whom English was the predominant language at home averaged 0.26 standard deviations higher than the other students). This is a moderate difference. The difference was a little larger for responding to art tasks (+.32) and negligible for art making task (-.01). There were statistically significant differences on 4 of the 17 tasks: *Art Objects* (p34), *Meeting House* (p36), *Wearable Arts* (p40), and *Link Task 6* (p43). Students for whom English was the predominant language spoken at home performed better on all four tasks. There were also differences on five questions of the *Year 4 Art Survey* (p46): Students whose predominant language at home was not English thought they had more opportunity at school to do drawing, collage, carving, model making, and work with fabrics or weaving.

For year 8 students, the mean effect size across the 17 tasks was +.26 (students for whom English was the predominant language at home averaged 0.26 standard deviations higher than the other students). This is a moderate difference. The difference was substantially larger for responding to art tasks (+.45) and slightly in the opposite direction for art making tasks (-.10). There were statistically significant differences on six of the 17 tasks (all responding to art tasks): Students for whom English was the predominant language spoken at home performed better on these 6 tasks. There was also a difference on two questions of the *Year 8 Art Survey* (p47): Students whose predominant language at home was not English thought they learned more in art at school and had more opportunity at school to do collage.