

5 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 (p8).

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 50 percent

of tasks at both year 4 and year 8. 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 four of the 45 tasks. Students from intermediate schools scored higher on the four tasks: *Population Change* (p13), *Playground Map* (p14), *Link Task 10* (p33) and *Link Task 19* (p44).



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 were differences among the subgroups on two of the 37 tasks: students from small schools scored highest on *Population Change* (p13) and *Link Task 18* (p44).

For year 8 students, there was a difference on just one of the 45 tasks. Students from small schools scored lowest on *Link Task 21* (p44).

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 four of the 37 tasks. Students from main centres scored lowest on *Room One Winter Sports* (p26), students from provincial centres scored highest on *Population Change* (p13) and students from rural areas scored lowest on *Campground* (p42) and *Link Task 18* (p44).

For year 8 students, there were differences among the three subgroups on two of the 45 tasks. Students from provincial towns scored lowest on *Link Task 15* (p33) and *Link Task 21* (p44).

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 two of the 37 tasks. Students from the rest of the North Island scored lowest on *Link Task 8* (p32) and *Link Task 16* (p44).

For year 8 students, there were differences among the three subgroups on two of the 45 tasks. Students from the North Island excluding Auckland scored lowest on *Link Task 15* (p33), with students from the South Island scoring lowest on *Link Task 21* (p44).

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 10 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 19 of the 37 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 19 tasks, with students in medium decile schools somewhere between.

For year 8 students, there were differences among the three subgroups on 33 of the 45 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 33 tasks, with students in medium decile schools generally closer to the students in high decile schools.

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, 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.

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 37 tasks was $.00$. In other words, on average there was no difference between males and females. There were statistically significant differences on two of the 37



tasks. Boys performed better on *Link Task 11* (p32), while girls performed better on *Movie Prices* (p28).

For year 8 students, the mean effect size across the 45 tasks was $-.08$ (girls averaged 0.08 standard deviations higher than boys). This is a small difference. There were statistically significant differences on five of the 45 tasks. Boys performed better on *Population Change* (p13). Girls performed better on the other four tasks: *Room One Winter Sports* (p26), *Favourite Fruits* (p41), *Link Task 19* (p44) and *Link Task 21* (p44).

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 37 tasks was $+0.33$ (Pakeha students averaged 0.33 standard deviations higher than Māori students). This is a moderate difference. There were statistically significant differences on 18 of the 37 tasks, with Pakeha students performing better on all 18 tasks.

For year 8 students, the mean effect size across the 45 tasks was $+0.40$ (Pakeha students averaged 0.40 standard deviations higher than Māori students). This is a moderate to large difference. There were statistically significant differences on 32 of the 45 tasks: Pakeha students performed better on these 32 tasks.

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 37 tasks was $+0.50$ (Pakeha students averaged 0.50 standard deviations higher than Pasifika students). This is a large difference. There were statistically significant differences on 22 of the 37 tasks: Pakeha students performed better on all 22 tasks.

For year 8 students, the mean effect size across the 45 tasks was $+0.70$ (Pakeha students averaged 0.70 standard deviations higher than Pasifika students). This is a large difference. There were statistically significant differences on 35 of the 45 tasks: Pakeha students performed better on all 35 tasks.

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 $+0.35$ (students for whom English was the predominant language at home averaged 0.35 standard deviations higher than the other students). This is a moderate difference. There were statistically significant differences on 16 of the 37 tasks: students for whom English was the predominant language spoken at home performed better on these 16 tasks.

For year 8 students, the mean effect size across the 45 tasks was $+0.27$ (students for whom English was the predominant language at home averaged 0.27 standard deviations higher than the other students). This is a moderate difference. There were statistically significant differences on 12 of the 45 tasks: students for whom English was the predominant language spoken at home performed better on these 12 tasks.