9 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 the total score for each task, created as described in Chapter 1 (p5).

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 for tasks reporting results for individual students 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 tasks administered to teams or groups of students, p = .05 was used as the critical level, to compensate for the smaller numbers of cases in the subgroups.

For the first four of the five school variables, statistically significant differences among the subgroups were found for less than 16 percent of the tasks at both year levels. For the remaining variable, statistically significant differences were found on nearly two thirds 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 Size

Results were compared from students in larger, medium sized, and small schools (exact definitions were given in Chapter 1).

For year 4 students, there were differences among the three subgroups on four of the 36 tasks. Students attending small schools scored lowest, and students from large schools highest on *We Need a Leader* (p15), *Link Task 7* (p30) and *World Current Events* (p40). Students from small schools scored highest on *Link Task 8* (p30). There

were no differences on questions of the *Social Studies Survey* (p49).

For year 8 students there were differences among the three subgroups on two of the 41 tasks. Students from medium size schools scored highest and students from small schools lowest on *Rodney's Window* (p38). Students from small schools scored highest on *Up and Down* (p47). There was also a difference on one question of the *Social Studies Survey* (p50), with students from medium sized schools reporting the fewest opportunities to learn about why people have different ideas (question 18).

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 among the three subgroups on three of the 36 tasks. Students from main centres scored highest on all three tasks: *Link Task 6* (p30), *Link Task 9* (p36), and *World Current Events* (p40). There were no differences on questions of the *Social Studies Survey* (p49).

For year 8 students, there was a difference among the three subgroups on one of the 41 tasks. Students from rural areas scored highest and students from main centres lowest on *Link Task 19* (p47). There were also differences on two questions of the *Social Studies Survey* (p50). Students from rural areas were most positive and students from provincial cities least positive about learning about other places in New Zealand and how people live there (question 8). Students from main centres reported most opportunities to learn about the way people work together and do things in groups (question 14), with students in provincial cities reporting the least opportunities.

School Type

Results were compared for year 8 students attending full primary, intermediate (or middle) schools and year 7 to 13 high schools.

In comparing students attending full primary and intermediate (or middle) schools, there were differences on two of the 41 tasks. Students attending intermediate (or middle) schools scored higher than students attending full primary schools on *Link Task 6* (p30), but the reverse was true on *Link Task 13* (p42). There were no differences on the questions of the *Social Studies Survey* (p50).

In comparing year 8 students attending intermediate (or middle) schools to those attending year 7 to 13 high schools, there were differences on two of 34 tasks. Students attending year 7 to 13 high schools scored higher than students attending intermediate (or middle) schools on both tasks: *Earthquake Disaster (Y8)* (p17) and *Coat of Arms* (p30). There was also a difference on one question of the *Social Studies Survey* (p50), with students from intermediate schools reporting more opportunities to learn about why people have different ideas (question 18).

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 36 tasks. Students from the South Island scored lowest and students from Auckland scored highest on *Link Task 6* (p30). Students from Auckland scored highest on *Kaiwakamoana* (p34). Students from the rest of the North Island (excluding Auckland) scored lowest on *Link Task 9* (p36). There was also a difference on one question of the *Social Studies Survey* (p49): students from Auckland were most positive and students from the South Island least positive about learning or doing more social studies as they got older (question 5).

For year 8 students, there were differences among the three subgroups on six of the 41 tasks. Students from the South Island scored highest and students from Auckland lowest on four tasks: *Link Task 8* (p30), *Relief Map* (p32), *New Zealand Places* (p33) and *Link Task 13* (p42). On the fifth task the opposite was true: students from Auckland scored highest and students from the South Island lowest on *Link Task 6* (p30). Students from the North Island other than Auckland scored highest on *Pōwhiri* (p23). There was also a difference on one question of the *Social Studies Survey* (p50): students from Auckland thought they learned the most in social studies at school (question 2).

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 resulting index takes into account household income levels and categories of employment. It 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 decile group, the middle four deciles (4-7) formed the medium decile group and the top three deciles (8-10) formed the high decile group. Results were compared for students attending schools in each of these three groups.

For year 4 students, there were differences among the three subgroups on 19 of the 36 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, they are not listed here. On *Link Task 6* (p30), which involved Māori knowledge, students in the low decile group scored highest, with students in the medium decile group lowest. For the other 18 tasks, performance was lowest for students

in the low decile group, often with quite a large gap to the students in the medium decile group. Students in the high decile group performed better than students in the medium decile group on most tasks, but these differences were generally quite small. There were significant differences on five questions of the *Social Studies Survey* (p49). Students in the low decile group were more positive than students in the high decile group on four questions: wanting to study more social studies at school (question 3), wanting to learn or do more social studies as they got older (question 5), enjoying learning about the way people work together and do things in groups (question 6), and enjoying learning about the work people do and how they make a living (question 9). Students in the low decile group also reported having more opportunities to learn about the work people do and how they make a living (question 17).

For year 8 students, there were differences among the three subgroups on 23 of the 41 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, the specific tasks are not listed here. In each case, performance was lowest for students in the low decile group, often with quite a large gap to the students in the medium decile group. Students in the high decile group performed better than students in the medium decile group on most tasks, but these differences were generally quite small. There were no differences among groups on the questions of the *Social Studies Survey* (p50).

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 cycle of the Project that took place from 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 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 +.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 effect-size procedures.

For year 4 students, the mean-effect size across the 29 tasks was .01 (girls averaged 0.01 standard deviations higher than boys); this is a negligible difference. There were statistically significant ($p < .01$) differences favouring boys on three of the 29 tasks, all involving factual geographic knowledge demonstrated on a laptop computer: *New Zealand Places* (p33), *Link Task 10* (p36) and *Link Task 11* (p36). There were no differences on questions of the *Social Studies Survey* (p49).

For year 8 students, the mean-effect size across the 34 tasks was .03 (girls averaged 0.03 standard deviations higher than boys); this is a very small difference. There were statistically significant differences on nine of the 34 tasks. Girls performed better than boys on five of the nine tasks: *Link Task 2* (p21), *Link Task 7* (p30), *Olivia* (p35), *Link Task 13* (p42) and *Manda* (p44). Boys performed better than girls on the other four tasks: *New Zealand Places* (p33), *Link Task 12* (p42), *Link Task 15* (p42) and *Up and Down* (p47). There were differences on three questions of

the *Social Studies Survey* (p50), with girls more positive than boys about doing social studies at school (question 1), liking to do more social studies at school (question 3) and wanting to learn or do more social studies as they got older (question 5).

Ethnicity

Results achieved by Māori, Pasifika and Pakeha (all other) students were compared using 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.

Pakeha-Māori Comparisons

For year 4 students, the mean-effect size across the 29 tasks was .24 (Pakeha students averaged 0.24 standard deviations higher than Māori students). This is a moderate difference. There were statistically significant differences (p <. 01) on 17 of the 29 tasks. Māori students scored higher than Pakeha students on two tasks involving Māori contexts: *Pōwhiri* (p23) and *Link Task 6* (p30). Pakeha students scored higher than Māori students on the remaining 15 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, they are not listed here. There were also differences on four questions of the *Social Studies Survey* (p49): Māori students were more positive than Pakeha students about wanting to study more social studies at school (question 3) and wanting to learn or do more social studies as they got older (question 5) and learning about the work people do and how they make a living (question 9). Māori students also reported having more opportunities to learn about the work people do and how they make a living (question 17).

For year 8 students, the results were similar. The mean-effect size across the 34 tasks was .24 (Pakeha students averaged 0.24 standard deviations higher than Māori students). This is a moderate difference. There were statistically significant differences on 14 of the 34 tasks. Māori students scored higher than Pakeha students on two tasks involving Māori contexts: *Pōwhiri* (p23) and *Link Task 6* (p30). Pakeha students scored higher than Māori students on the remaining 12 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, they are not listed here. There was also a difference on one question of the *Social Studies Survey* (p50): Māori students were more positive than Pakeha students about learning about how people lived in the "olden days" (question 12).

Pakeha-Pasifika Comparisons

Readers should note that only 28 to 42 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. Because of the relatively small numbers of Pasifika students, p = .05 has been used here as the critical level for statistical significance.

For year 4 students, the mean-effect size across the 29 tasks was .24 (Pakeha students averaged 0.24 standard deviations higher than Pasifika students). This is a moderate difference. There were statistically significant differences on 12 of the 29 tasks. Pasifika students scored higher than Pakeha students on one task involving a Māori context: *Pōwhiri* (p23). Pakeha students scored higher than Pasifika students on the remaining 10 tasks, spread fairly evenly across Chapters 3 to 7 (only one in Chapter 3). Because of the number of tasks involved, they are not listed here. There were also differences on six questions of the *Social Studies Survey* (p49). Pasifika students were more positive than Pakeha students about doing social studies at school (question 1), wanting to study more social studies at school (question 3) and wanting to learn or do more social studies as they got older (question 5). Pasifika students also reported that their class more often did really good things in social studies (question 4), and that they experienced more opportunities to learn about the way people work together and do things in groups (question 14) and how people lived in the "olden days" (question 20).

For year 8 students, the mean-effect size across the 34 tasks was 0.42 (Pakeha students averaged 0.42 standard deviations higher than Pasifika students). This is a large difference. There were statistically significant differences on 27 of the 34 tasks. Pasifika students scored higher than Pakeha students on two tasks involving Māori contexts: *Pōwhiri* (p23) and *Link Task 6* (p30). Pakeha students scored higher than Pasifika students on the remaining 25 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, they are not listed here. There were also differences on five questions of the *Social Studies Survey* (p50). Pasifika students were more positive than Pakeha students about wanting to study more social studies at school (question 3). Pasifika students also reported that they experienced

more opportunities to learn about other places in New Zealand and how people live there (question 16), the work people do and how they make a living (question 17), why people have different ideas (question 18) and living in the future (question 21).

Home Language

Results achieved by students who reported that English was the predominant language spoken at home were compared, using effect-size procedures, with the results of students who reported predominant use of another language at home (most commonly an Asian or Pasifika language). Because of the relatively small numbers in the "other language" group, p = .05 has been used here as the critical level for statistical significance.

For year 4 students, the mean-effect size across the 29 tasks was .08 (students for whom English was the predominant language at home averaged 0.08 standard deviations higher than the other students). This is a small difference. There were statistically significant differences on six of the 29 tasks. Students for whom English was the predominant language at home performed significantly better than the students who reported using another language at home on five of the tasks: *Saikoloni* (p14), *Link Task 1* (p21), *Link Task 8* (p30), *Relief Map* (p32) and *Link Task 11* (p36). The converse was true on *Olivia* (p35). There were also differences on four questions of the *Social Studies Survey* (p49). Students for whom the predominant language at home was not English were more positive than their English language counterparts about wanting to study more social studies at school (question 3), wanting to learn or do more social studies as they got older (question 5), learning about other places in the world and how people live there (question 7) and learning about the work people do and how they make a living (question 9).

For year 8 students, the mean-effect size across the 34 tasks was 0.23 (students for whom English was the predominant language at home averaged 0.23 standard deviations higher than the other students). This is a moderate difference. There were statistically significant differences on 14 of the 34 tasks. Students for whom English was the predominant language spoken at home scored

higher on all 14 tasks, spread evenly across Chapters 3 to 7. Because of the number of tasks involved, they are not listed here. There were also differences on six questions of the *Social Studies Survey* (p50). Students for whom the predominant language at home was not English reported that they experienced more opportunities in social studies at school to learn about the way people work together and do things in groups (question 14), other places in New Zealand and how people live there (question 16), why people have different ideas (question 18), what is happening now, in New Zealand and other countries (question 19), how people lived in the "olden days" (question 20) and living in the future (question 21).

Summary, with Comparisons to Previous Social Studies Assessments

Community size, school size, and school type (full primary, intermediate, or year 7 to 13 high school), and geographic zone did not seem to be important factors predicting achievement on the social studies tasks. The same was true for the 2001 and 1997 assessments. However, there were statistically significant differences in the performance of students from low, medium and high decile schools on 53 percent of the tasks at year 4 level (compared to 67 percent in 2001 and 53 percent in 1997), and 56 percent of the tasks at year 8 level (compared to 49 percent in 2001 and 73 percent in 1997).

For the comparisons of boys with girls, Pakeha with Māori, Pakeha with Pasifika students, and students for whom the predominant language at home was English with those for whom it was not, effect sizes were used. Effect size is the difference in mean (average) performance of the two groups, divided by the pooled standard deviation of the scores on the particular task. For this summary, these effect sizes were averaged across all tasks.

Year 4 girls averaged very slightly higher than boys, with a mean effect size of 0.01 (in 2001, year 4 boys had a small advantage with a mean effect size of 0.06). Year 8 girls averaged very slightly higher than boys, with a mean effect size of 0.03 (very similar to the mean effect size of 0.02 in 2001). As was also true in 2001, the *Social Studies Survey* (p50) results showed some evidence that year 8 girls were more positive than boys about social studies activities.

Pakeha students averaged moderately higher than Māori students, with mean effect sizes of 0.24 for both year 4 and year 8 students (the corresponding figures in 2001 were 0.28 and 0.32). Māori students were more positive than Pakeha students on four questions of the *Social Studies Survey* (p49) at year 4 level and one question at year 8 level.

Year 4 Pakeha students averaged moderately higher than Pasifika students, with a mean effect size of 0.24 (a noteworthy reduction in disparity from 0.47 in 2001). Year 8 Pakeha students averaged substantially higher than Pasifika students, with a large mean effect size of 0.42 (reduced from 0.51 in 2001). Pasifika students were more positive than Pakeha students on some questions of the *Social Studies Survey* (pp49-50) at both year levels.

Compared to students for whom the predominant language at home was English, students from homes where other languages predominated averaged slightly lower at year 4 level (mean effect size of 0.08) and moderately lower at year 8 level (mean effect size of 0.23). Comparative figures are not available for the assessments in 2001. Year 4 students whose predominant language at home was not English were more positive than their English language counterparts on some questions of the *Social Studies Survey* (p49).