

# Group Assessment

## Does gender composition of the group matter?

*Grace Grima*

This article provides a summary of the work presented in the monograph: **Grima, G. (2000) *Group Assessment: exploring the influences of the group gender composition*. New Zealand: Ministry of Education.**

“The discourse that continues to suggest that girls are disadvantaged in mixed gender settings needs to be questioned”

### Overview

In this article I give an outline of the National Education Monitoring Project and describe the position of the probe study within it. I then explain the design and method of the probe study. I also provide highlights from the results that directly answer the research questions set out. Finally, I discuss the implications of the findings of this study for educational practice.

### New Zealand's National Education Monitoring Project

New Zealand's National Education Monitoring Project (NEMP) commenced in 1993 with the task of assessing and reporting on the achievement of New Zealand's school children in all areas of the school curriculum. Children are assessed at two class levels: at Year Four (ages 8-9) and Year Eight (ages 12-13). Different curriculum areas are assessed each year, over a four-year cycle. This four year cycle also incorporates the assessment of attitudes and several social skills.

The information for this national project that provides a snapshot of children's educational achievements at the two class levels is gathered from random samples of students chosen annually. The annual national samples containing 1,440 students at each of the two levels represent about three percent of the children at those levels in New Zealand schools. The students selected at each level are divided into three subgroups (A, B and C). The subgroups, which consist of 120 four-member groups, attempt different tasks, many of which are repeated at the two age levels.

The assessments take place over two five-week periods between August and November and are conducted by experienced teachers, who are selected from a national pool of applicants to carry out the assessments for one period. The teachers attend a week of specialist training led by the NEMP staff. The training takes place just before the assessment period commences.

The teachers then work in pairs throughout the assessment period. In each school the assessments are spread over one week, and each student participates in about four hours of assessment activities during that week.

The assessment activities include:

- one-to-one tasks, where each student works individually with a teacher.
- station tasks, where four students work independently, moving around a series of stations where tasks have been set up.
- group tasks, where four students work collaboratively on the same task.

Participation in the one-to-one and the group tasks is recorded on videotape for subsequent analysis of both processes and products.

### The Case for Group Assessment

NEMP is one of the few large assessment projects that has taken up the challenge of assessing what children can do in groups as well as individually. The literature suggests a number of reasons for incorporating group assessment into performance assessment settings. These include the need to:

- link assessment more closely to the growing emphasis on small-group collaboration and co-operation in classroom instruction (Linn, 1993; Webb, 1995)
- send out a message to educators about the importance of group collaboration in classroom instruction (Wise & Behuniak, 1993, cited in Webb, 1993).
- provide authentic assessment that involves complex problems in realistic conditions (Crooks & Flockton, 1994; Webb, 1993)
- provide a milieu within which to measure interpersonal skills that relate to the social goals of education (Crooks & Flockton, 1994; Webb, 1993, 1997)
- make it possible to assess group productivity and effectiveness (Webb, 1993, 1997).

It was still unclear up to this point whether any type of group composition was advantageous over others in situations involving true group tasks (i.e. those tasks that require the group members to make use of their particular skills and resources in order to complete the tasks). Webb (1995) suggested that "certain groups may be unfair if they do not give students equal access to favourable group processes" (249). She advocated the need for research that "explores the effects of varying group compositions on processes and outcomes of assessments for different types of tasks to determine when and whether the group composition is a source of bias (225). My probe study responded to that need.

### The Probe Study

Given that NEMP involves a relatively large number of groups working on the same tasks and that the process is being video-recorded, the probe study provided a valuable opportunity to analyse the extent to which gender composition of a group influences the group's experience. Although the probe study focuses on the gender composition of groups, it acknowledges that the group experience may be influenced by several other structural factors, such as group size, student characteristics (e.g. age, personality, ability and ethnicity) and group ability composition (i.e. whether the groups are homogeneous or include students with a combination of two or three ability levels).

The study evaluated the extent to which groups with different gender compositions - four boys (4b), 3 boys 1 girl (3b1g), 2 boys 2 girls (2b2g), 1 boy 3 girls (1b3g) and 4 girls (4g) - provided boys and



**GRACE GRIMA**  
B.Ed.(Hons.) M.Ed.(Otago)  
Ph.D.(Otago) is a  
lecturer at the Faculty of  
Education and a Principal  
Research and Development  
Officer in Assessment at  
Matsec.



girls with a similarly productive and enjoyable experience during three tasks from different curriculum areas. The analyses focused on these dependent variables: individual participation, group interaction, co-operation and conflict as well as the group products. The independent variables were the five types of group gender compositions, the two age levels of the students and the different nature of the three tasks.

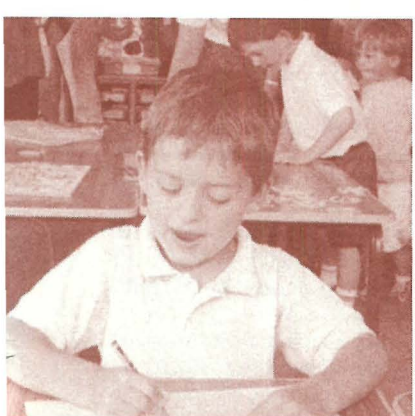
The tasks analysed in this study were the following NEMP tasks:

- A science task called *Separating Mixtures* where the children first discussed and then carried out an experiment, which involved separating a mixture, made up of flour, paperclips, polystyrene balls and ball bearings.

- A language task called *Question Time* where the children first brainstormed questions they could ask police officers and then chose the six that they would ask as a group and justified their choices.

- A technology task called *Space Game* where the children first played a board game called Planet Fero and then discussed ways of improving and marketing the game in two separate discussions. In each discussion, they chose and reported their best four ideas

The study also investigated how New Zealand children felt about working in groups with different gender compositions by means of a questionnaire. A post-task evaluation was also carried out in order to compare the children's views on one particular NEMP experience in the different group types. Finally, a small number of children were interviewed to gain further insight into their perspectives on group work and explore their interpretations of the events that occurred during one group task.



## The Research Questions

The probe study addressed the following questions:

1. Does the task involvement of boys and girls differ according to the gender composition of the group? Does it change with the children's age level and the nature of the tasks?
2. What is the relationship between the group gender composition and the levels of interaction, co-operation and conflict within groups? Do these levels remain consistent across the two age levels and across the different tasks?
3. What is the relationship between the group gender composition and task achievement? Does this relationship change with the children's age level and the nature of the tasks?
4. How do Year Four and Year Eight children feel about working in groups with different gender compositions?
5. Do the children's evaluations of their NEMP group experience vary across group types? Are these differences affected by the nature of the tasks?

## Research Methods

The research methods used in the study included videotaped recordings, a questionnaire and interviews as well as the groups' achievement records.

### Videotaped recordings

The choice of techniques for analysing group processes includes categories, checklists, rating scales and narrative accounts. I did some trial work using both categories and checklists but found both techniques unsatisfactory. Consequently I developed a structured observation schedule using rating scales and narrative accounts. As tools, the narrative accounts and the rating scales complemented each other because the former needed to be recorded while viewing the tape whereas the latter needed to be completed after the viewing.

### The video analysis

I originally planned to have, for each of the three tasks, ten groups of each of the five group types (4b, 3b1g, 2b2g, 1b3g, 4g) at both Year Four and Year Eight. However, as a result of the random sampling technique used in NEMP, it was rare to have ten same gender groups of boys and of girls. Therefore, when the number of groups available was less than ten, I included all the groups. When there were more than ten groups, I used systematic sampling to make my selection.

A computerised database for each of the three observation schedules was created using Filemaker Pro (Claris, 1995) and Cvideo (Envisionology, 1993). The videos were analysed in the order of the NEMP classification numbers rather than according to their group gender composition, thus reducing the possibility of developing a bias in relation to particular group types.

For each task, I chose video exemplars and developed descriptors for the levels of the rating scales during the initial viewing and then checked the coding of the sample of tapes during a second viewing. I viewed the tapes twice more, once starting with the Year Four tapes and once starting with the Year Eight tapes to ensure that I was as consistent

as possible in similarly coding the tapes at the two age levels. In total, I spent, 1,100 hours conducting the video analysis.

### The questionnaires

In order to explore how children feel about working in groups with different gender compositions generally and how they felt during one of the NEMP group experiences in particular, I used a questionnaire with all the students who participated in NEMP in 1996. This tool served as a self-report for the children. With the older age group the questionnaire was self-administered (i.e. the students read the questions and marked their answers). For the younger age group, one teacher administrator read the questions (and the answers when these included words) to a group of four children. The children then marked their own answers on their sheet. There was a hundred percent return of the questionnaires that had been completed (i.e. 94% of the national sample participating in NEMP in 1996). In total, 2,716 questionnaires were analysed.

### The interviews

Working within the time constraints of a national project, my schedule allowed me to interview the 23 eight-year-olds that were randomly selected to participate in the NEMP assessments in the Dunedin city area in 1996. I interviewed the children in a one-off situation on the school premises at times during the assessment week when they were not involved in other activities relating to NEMP. I used two interviewing techniques: the structured interview and a type of focused interview called stimulated recall. I found both techniques to be appropriate in a one-off interview situation. In this article however, I only present responses from the segment of the structured interview that relate to the final research question that is addressed.

### The achievement records

The achievement records were used to analyse the products of groups with different gender compositions. Marking schedules for the three tasks were developed by NEMP and used by the teacher markers who were involved in marking the tasks.

## Highlights from Results

### Participation within the five group types

- Overall, the video analysis did not identify any group types that had higher participation levels consistently across the three tasks.

- Especially at Year Four, there was a tendency for the minority student in the 3b1g and the 1b3g group types to participate less than the other group members and/or to participate less than members of their gender group working in other group settings. At the same time, the minority student tended to become highly involved in the organisation aspect of the group task.

- At Year Eight, both boys and girls, participated less in the 2b2g groups than in the other group types.

- There was no indication at either Year Four or Year Eight that girls were disadvantaged in the mixed gender groups compared to boys. And there was no evidence of domination by boys in the mixed gender groups across the three tasks.



## Group processes within the five group types

### Interaction

● Overall, across the two age groups, the 4b groups were observed to interact the most.

### Co-operation

● The older groups co-operated more than the younger ones in all of the group types.

● The 4g groups had the highest co-operation mean at both Year Four and Year Eight. However, there was no indication that the same gender groups co-operated more than the mixed gender groups.

● The 3b1g groups stood out as the groups that experienced most difficulty working together in several activities at Year Four.

### Conflict

● Conflict was more common in the Year Four groups than in the Year Eight groups.

● Overall, conflict was present almost evenly in all of the group types at both age levels.

● In general, the 4g groups tended to get highly involved in conflict during verbal activities whereas the 4b groups tended to become highly involved in conflict in activities when they were doing something physical (carrying out an experiment, playing a board game).

Overall, the analysis of the relationship between the group gender composition and levels of interaction, co-operation and conflict showed that the experiences were relatively similar across the different group types. Certainly, there were no clear divisions between the experience in the same gender and the mixed gender groups. In general, it was one group type that stood out in the different analyses.

## Joint products of the five group types

● At Year Four, achievement on the group product was lower in the same gender groups than in the mixed gender groups.

● In the Year Eight, there was no pattern that clearly separated out the same gender and the mixed gender groups.

● Only the 1b3g groups scored higher than the average mean score at both ages.

● Achievement in the different group types appeared to be influenced at times by the **subject area** (e.g. the 4g groups scored higher than the average

mean score on the language task at both age groups), **the task content** (e.g. at Year Four the 4g groups scored higher than the average mean score in the language task and lower than the average mean score in both of the discussions of the technology task) and **the age group** (e.g. the 4g groups scored higher than the average mean score in the experiment of the science task at Year Four and lower than the average mean score in the same activity at Year Eight).

● Analyses of the relationship between group achievement and interaction, co-operation and conflict did not yield consistent results across the three tasks and at the two age levels. The most common pattern found was a negative relationship between group conflict and achievement.

## Children's perspectives on group gender composition

● At both ages, boys and girls generally felt least positive about working in a group with members of the other gender only. The children talked about discomfort, insecurity, peer pressure and a lack of mixed gender friendships.

● At Year Four, both boys and girls enjoyed working in same gender groups the most. Boys as well as girls talked about fraternity and collegiality amongst members of their gender group:

● At Year Eight, boys and girls responded equally favourably to same gender and mixed gender groups with two boys and two girls.

● At both ages, girls responded more favourably than boys did to the idea of working in mixed gender groups with equal numbers of boys and girls.

## Children's post-task evaluations

● Both boys and girls, at both ages, liked their NEMP group placement the least in the group type where they were outnumbered.

● Overall, boys in the 1b3g groups seemed less positive about their experience than boys who were placed in the other group types.

● Boys enjoyed their experience in same gender groups more than girls.

● Girls placed in same gender groups were not more positive about their group experience than those placed in mixed gender groups.

## Implications for Practice

● Overall, the study showed that the gender composition of a group is not a salient factor in children's task groups where they have the opportunity to focus on a shared goal. The goal became an important uniting factor, which overrides and reduces the salience of gender issues. This study concludes that individual differences between children are more important than gender in determining their contribution to and participation in group work.

● The study also shows that the discourse that continues to suggest that girls are disadvantaged in mixed gender settings needs to be questioned. On many occasions I found that it was boys who were left out or sidelined in mixed gender groups. Teachers should not assume that this does not happen and neither should they assume that boys always dominate in mixed gender settings. Overall, the

findings of this study challenge the almost stereotypical belief that boys and girls cannot work together effectively.

● This study shows that having equal numbers of boys and girls in a group does not necessarily result in an equitable experience, nor does this structure protect group members (boys as well as girls) from becoming sidelined in the activity.

● Moreover, this study also shows that children's experiences in same gender groups are not necessarily more equitable than in the mixed gender groups, nor are they consistently more productive, especially for girls. This finding strongly challenges the belief that the behavioural problems observed in mixed gender groups in particular, and co-educational schooling in general, cease to exist when boys and girls are placed in separate groups and/or settings

● Despite the finding that group gender composition had little effect on children's behaviour and achievement in small groups, there were differences in the children's attitudes towards working in groups with different gender compositions. The questionnaire and interview data showed that at both age levels a large number of children perceived the outnumbered student in a gender-imbalanced group to be disadvantaged. Although the data obtained from observing the different tasks at the two age levels did not consistently back up this perception, teachers need to be careful that they do not prematurely and regularly place students in a group situation where they are outnumbered. This does not mean that teachers should always avoid such placements, but rather that they should implement them gradually and carefully and provide the children with adequate support. Moreover, these placements should be complemented with experiences in other settings in which children initially feel more comfortable (i.e., same-gender and gender-balanced groups). The goal is to ensure that children have opportunities to develop a positive attitude towards all forms of group work and the necessary skills to function effectively when they are the only boy or girl in a group. This gradual process, well supported, should enable children to get accustomed to working in all possible group situations.

## References

- Claris, C. (1995). FileMaker Pro (Version 3.0). California: Claris.
- Crooks, T & Flockton, L. (1994). National Education Monitoring Project – Phase II. Dunedin: Educational Assessment Research Unit, University of Otago.
- Envisionology (1993). Cvideo (Version 1.5). San Francisco: Knowledge Revolution.
- Linn, R. (1993). Educational Assessment: expanded expectations and challenges. *Educational Evaluation and Policy Analysis*, 15(1), 1-16.
- Webb, N. M. (1993). Collaborative group versus individual assessment in Mathematics: processes and outcomes. *Educational Assessment*, 1(2), 131-152.
- Webb, N. M. (1995). Group collaboration in assessment: multiple objectives, processes and outcomes. *Educational Evaluation and Policy Analysis*, 17(2), 239-261.
- Webb, N. M. (1997). Assessing students in small collaborative groups. *Theory into Practice*, 36(4), 205-213.