Group Assessment: exploring the influences of the group gender composition

Grace Grima

2002

DEDICATION

To my mother, Maria Grima.

ACKNOWLEDGMENTS

I initially wish to acknowledge the New Zealand Vice Chancellor's Committee for providing me with a Commonwealth scholarship and thus making it possible for me to return to New Zealand to conduct my PhD study at the University of Otago. I also acknowledge the University of Malta for providing me with leave during this period. Most of all, I acknowledge the Educational Assessment Research Unit of the University of Otago and its co-directors Terry Crooks and Lester Flockton for providing me with the wonderful opportunity to carry out my PhD project as a probe study on the National Educational Monitoring Project.

My deepest gratitude goes to my supervisors Terry Crooks and Anne Smith. Throughout this long journey, they were always there for me, providing me with professional support and guidance as well as with the genuine friendship that made my study such a worthwhile experience. My deepest thanks go to the children who participated in the National Educational Monitoring Project in 1996 and 1997. They made this project possible.

I would also like to acknowledge all the staff members at the Educational Assessment Research Unit and the Children's Issues Centre at the University of Otago who provided me with support in numerous ways. My sincere thanks go to Terry, Lester, Janice, Linda, Liz, Miriam, Nicole, Robyn, Janet, Lee and Lizzie, as well as to Anne, Nicki, Rachael, Michael, Megan, Karen and Joanna.

Particular thanks go to my special friends who were there to listen when I needed to talk, who provided me with companionship when I got lonely and who have enriched my life with their friendship. Janice and Mark, Anne and John, Ros and Ian, Terry and Mina, Linda and Clive, Carmen and Richard, Clarissa and Jean Paul, and Patria and Johnson, thank you. I would also like to thank my younger friends who have a special place in my heart: Francesca, George and Peter, who live in New Zealand, and Ella and Edward, who live in Malta. Thank you all for being my friends.

Last but not least I would like to thank my family, Guzeppi, Maria and Frida, for supporting my dreams. Grazzi minn qalbi.

CONTENTS

TABLES

ABSTRACT

This multi-method study explored the extent to which children's task groups, each containing four members and with different gender compositions, provided their members with a productive and enjoyable experience. The study included three tasks from different curriculum areas (science, language and technology) completed at two age levels (year 4 and year 8). For each task, approximately 90 groups were drawn randomly from the larger samples participating in the New Zealand National Education Monitoring Project (NEMP). The analysis of their work focused on a number of variables: the group members' individual participation levels; several group processes (interaction, co-operation and conflict); and the group products.

The study also investigated how New Zealand children at years 4 and 8 felt about working in groups with different gender compositions. A post-task evaluation of one of the NEMP tasks was also carried out in order to compare the children's views on their experiences in the different group types.

The experience in groups with four boys (4b), three boys and one girl (3b1g), two boys and two girls (2b2g), one boy and three girls (1b3g) and four girls (4g) did not vary consistently across tasks and age levels. No group types were consistently observed to stand out in the analyses at either year 4 or year 8. However, at year 4, there was a tendency for the minority student, especially in the 1b3g groups, to participate less than the other group members and/or to participate less than those members of their gender group who were working in other group settings.

At year 4 the children preferred same-gender groups, but at year 8 they responded equally favourably to participating in same-gender and balanced mixed-gender groups. The post-task evaluations showed that the experience was less enjoyable and less productive for the minority student in the 3b1g and the 1b3g groups.

This empirical study did not emerge from any one theoretical paradigm. A variety of theories (expectation states theory, social role theory, structural numerical proportions theory, group cognition theory and post-structuralist theory) informed the study and, in turn, were used to interpret the results.

Previous research on group dynamics during tasks has often observed stages of tasks or activities that formed part of tasks and used these observations as the basis for generalised conclusions about the task as a whole or even about group tasks in general. In this study, video analysis showed that group dynamics were, in fact, inconsistent across the activities that made up each task. An accurate account of group processes occurring in different tasks only emerged when the different activities comprising these tasks were analysed separately. ■

1

GROUP WORK IN EDUCATION

The aim of this introductory chapter is to provide an overview of the literature on group work in education that contributes to the framework of this empirical study. I first outline the variety of instructional strategies that involve groups in education and identify the features of group work. I then discuss the rationale for this classroom activity and list the benefits that have been attributed to it. I conclude by examining a variety of structural factors that have been reported to influence the experience of group work.

DEFINITIONS OF GROUP WORK

Various instructional strategies utilised within a classroom context involve small groups. Galton and Williamson (1992, 10) have classified these as follows:

TYPE	TASK DEMAND	OUTCOME
Seating groups	Students have separate tasks	Different outcomes; each student completes a different assignment
Working groups	Students have the same task	Same outcome; students complete the same assignment independently
Co-operative group	Each student has a separate but related task	Joint outcome; each student has a different assignment
Collaborative group	Students have the same task	Joint outcome; all students share same assignment.

Not all of these groups involve group work because this process is distinguished by two features present in Cohen's (1986, 1-2) definition: 'students *work together* in a group small enough so that everyone can participate on a task that has been clearly assigned ... Moreover, students are expected to carry out the task *without the direct and immediate supervision of the teacher*' [emphasis mine]. Therefore, situations in which children sit in groups but do not work as groups (e.g., sitting and working groups) do not involve group work because (a) the students are not engaged in a common task; (b) they are not interdependent in the performance of that task; and (c) they do not need to interact in its pursuit (Webb & Palinscar, 1996).

Although distinctions between co-operative and collaborative group work are not always made explicit in the literature (e.g. that on co-operative learning), the group processes that evolve in these two situations are not identical. In co-operative group work, 'pupils work on the same task but all have individual assignments which eventually are put together to form a joint outcome' (Galton & Williamson, 1992, 10). However, in collaborative group work all members of the group operate together on all aspects of the task and contribute to a single outcome. Collaborative group work generally refers to three or more children working

together, whereas peer collaboration refers to a pair of children working together on the same task (Damon & Phelps, 1989). Although the dynamics of the two situations are not completely identical because of the different number of students involved in the activity, several of the processes that occur in the learning context of peer collaboration are also applicable in the context of collaborative group work.

In the literature, therefore, the term group work refers to two activities that involve small groups. Co-operative group work involves a division of labour, while collaborative group work involves all of the students working together on all aspects of the task.

GROUP WORK — A RATIONALE

Group work has been the focus of much research in the last three decades. This interest in children working together in small groups has been attributed to an increased awareness of the link between learning and social interaction. As Bennett (1994, 50–51) observes,

> there is a realisation among educators of the value of interpersonal processes in learning and social relationships, an increasing awareness of the value of co-operation and problem-solving in the development of understanding and a desire to move away from instructional models which view teachers as the only source of knowledge and skills.

This realisation is related, in part, to the constructivist view of the learner. Webb and Palinscar (1996, 844) explain that

> as instructional theorists turn their attention to contextualized practice, there is heightened interest in situations where elaboration, interpretation, explanation and argumentation are integral to the activity of the group and to where learning is supported by other individuals. In this sense, constructivism holds that cognition is an outcome of social processes.

The theories of Piaget and Vygotsky both refer to the effects of the social context on individuals' cognitive growth (Tudge & Rogoff, 1989). Although Piaget was primarily concerned with individual development, he believed that discussion between children had a role to play in cognitive development. Vygotsky's theory places a more central focus on social interaction as a medium in which children learn and develop.

From a Piagetian perspective, children gain social and cognitive benefits from peer interaction (Smith, 1998). The social benefits include improved communication skills and an increased awareness of the perspectives of others. The cognitive benefits come from the children's motivation to re-examine their own conceptions in the light of the perspectives of others. Piaget (1967, 163) commented that 'without interchange of thought and co-operation with others the individual would never come to group his operations into a coherent whole'. Thomas (1994, 8) suggests that, from this perspective,

> peer interaction and social experiences in general derive their importance from the influence they exert on equilibration through the existence of cognitive conflict... [which] is a perceived sense of contradiction between what the child believes and feedback the child receives on those beliefs. If the child is aware of the contradiction, the experience has a disequilibrating effect which encourages the child to construct new conceptions that fit better with the feedback that she is receiving.

Cognitive conflict hence acts as a catalyst for change. 'Such interaction between peers, Piaget argued, leads children to reconsider their ideas' (Rogoff, 1990, 147). According to Webb and Palinscar (1996), Piaget regarded social exchanges between children and adults as unlikely to lead to the kinds of cognitive development that exchanges with other children promote. Damon (1984) argues that giving up current understandings to reach a new perspective is best attained by an exchange

of ideas on an equal basis. He suggests a number of reasons why peers act as an effective source of cognitive conflict: the language spoken can be understood; children tend to speak directly and openly; they tend to take feedback from other children seriously; and they find the situation is less emotionally threatening than is receiving corrective feedback from an adult.

Webb and Palinscar (1996, 845) conclude that 'social interaction is considered from the perspective of how effective it is in creating conflict within the individual ... we begin by considering the individual and then move to the social interaction.'

From a Vygotskian perspective, however, social interaction is considered primary, because all higher mental functions are seen to develop through interaction either with adults or with peers. In this view, there is a gap between what children can do in conjunction with others and what they can do alone (Galton & Williamson, 1992). As Vygotsky (1962) said, 'what a child can do today in co-operation, tomorrow he will be able to do on his own'. Without interaction with others, children cannot internalise new skills, and it is only after they have internalised these skills that they can carry them out independently. Vygotsky (1978, 90) therefore was stressing the social nature of learning and development when he stated that

> learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in his environment and in co-operation with his peers.

This perspective highlights the importance of social interaction in learning and emphasises in particular the role of negotiation and sharing in the classroom (Bennett, 1994). As Slavin (1987, 1162) observes,

> collaborative activity among children promotes growth because children of similar ages are likely to be operating within one another's zones of proximal development,[1] modelling in the collaborating group behaviours more advanced than they could perform as individuals.

This notion constitutes an attack on the view of learning which assumes that intellectual competence is a result of a child's largely unassisted activities (Wood, 1987).

In New Zealand, the *National Curriculum Framework* recognises group work both as a process through which important skills can be learned and as a skill that students need to develop in order to function effectively in society (Ministry of Education, 1993). This document lists eight groupings of essential skills[2] that are considered 'important for students to achieve their potential and to participate fully in society' and stresses that a number of these skills 'may be developed through group activities. Furthermore, many of these skills will enable individuals to operate more effectively in group situations' (17). Bossert (1989, 225) summarises the perspective of many advocates of group work when he says that 'in the classroom, co-operation is both a skill necessary for the accomplishment of learning activities and a general norm to be learned'.

---

[1] Vygotsky's (1978, 86) notion of the zone of proximal development (ZPD) is defined as 'the distance between the actual development level as determined by independent problem-solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers'.

[2] These are (i) communication, (ii) numeracy, (iii) information, (iv) problem-solving, (v) self-management, (vi) competitive work and study, (vii) social and co-operative work and study and (viii) physical work and study.

POTENTIAL BENEFITS OF GROUP WORK

Small group contexts that allow students to work on learning tasks together provide the potential for numerous cognitive and socio-emotional benefits, as noted in the literature on co-operative learning and group problem-solving in the classroom, and in the theoretical perspectives on learning discussed above. According to Yeomans (1983, 100),

> the main claim made on behalf of group work is that it is the very process of group work itself that makes it a valuable experience. Today more consideration is placed on the processes involved in learning and the quality of the learning experience. Advocates of group work believe the processes involved in collaboration enable children of all abilities to learn more. This is achieved in a number of ways which interrelate and reinforce each other.

COGNITIVE BENEFITS

1 Group work makes it possible for students to have an active role in learning (Sharan, 1990). Active involvement has been reported to increase students' time on task (Cohen, 1986; Slavin, 1990a) and to decrease disruptive behaviour stemming from being bored (Sharan, 1990).

2 Group work also affords the externalisation of thought processes, social facilitation and socially monitored attentiveness to the task. These processes have the potential to lead to better learning and transfer of knowledge (Salomon & Globerson, 1989).

3 Working together and talking things through help children explore and handle new ideas (Plowden, 1967, cited in Yeomans, 1983) and master cognitive processes such as verification and criticism (Damon, 1984).

4 Through mutual feedback and debate, peers motivate one another to abandon misconceptions and search for better solutions (Damon, 1984). In addition, they are able to build on one another's ideas to construct the knowledge, skills and understanding that they did not have before group work (Damon & Phelps, 1989; Marland, 1977, cited in Yeomans, 1983).

5 Interaction with others may produce discrepancies between a child's view and new information, giving rise to cognitive conflict. One way that internal cognitive conflict can be manifested in interaction with others is overt conflict. This encourages individuals to explain and justify their positions, raises uncertainties about their beliefs, encourages them to seek new information and helps them understand alternative points of view, all of which can promote learning (Brown & Palinscar 1989, cited in Webb, 1994; Johnson & Johnson, 1979).

6 The group context provides students with the opportunities to learn, internalise and use problem-solving strategies that other students use or that are created jointly with others (Bearison, Magzamen & Filardo, 1986, Bossert, 1989, and Brown & Palinscar, 1989, all cited in Webb, 1995). Therefore, the combined interaction of the group has the potential to generate more learning than would be achieved by children individually (Maier, 1970, cited in Yeomans, 1983). Knight and Bohlmeyer (1990), cited in Webb (1993), explain that collaborative group performance usually exceeds individual performance because of the cognitive factors (e.g., more intellectual resources) and social factors (e.g., increased task motivation) available.

7 Grouping students provides them with the opportunity to help one another learn (Webb, 1995). Fellow students can be particularly good sources of help because they may understand better than the teacher what other students do not understand.

Also, they can direct the attention of other students to the relevant features of a problem they do not understand and can explain concepts in familiar terms (Brown & Palinscar, 1989, Noddings, 1985, and Vedder, 1985, all cited in Webb, 1997).

8 The opportunity to give explanations to others, especially when the material is complex and requires integration and reorganisation, is beneficial for understanding and applying ideas. Putting concepts into words (within the context of explaining to a peer) is helpful for concept attainment (Durling & Shick, 1986, cited in Cohen, 1986). The process whereby students give explanations either to help others or defend their own ideas provides them with the opportunity to clarify and reorganise material in new ways, to recognise and fill in gaps in understanding and to construct more elaborate conceptualisations than they would when learning material by themselves (King, 1992, and Yackel, Cobb & Wood, 1991, both cited in Webb, 1995).

9 Conversely, by receiving explanations, students fill in gaps in their understanding, correct any misconceptions and strengthen connections between new information and previous learning (Mayer, 1984, and Wittrock, 1990, both cited in Webb, 1994).

10 Furthermore, the experience of group work provides students with opportunities for active practice in oral communication (Cohen, 1986). In the group situation, students are forced to use their own language to express their ideas (Barnes & Todd, 1977), and they need to clarify the meaning to themselves and the group (Yeomans, 1983). Hence they are more likely to use strategies that involve higher-level reasoning and that avoid errors in reasoning (Johnson & Johnson, 1985b, cited in Nastasi & Clements, 1991). Such skills enable students to develop deeper learning and become more autonomous learners (Murray, 1988).

SOCIO-EMOTIONAL BENEFITS

1 Working in groups helps children develop self-confidence and independence, which increases their feelings of self-esteem (Sharan, 1980, cited in Yeomans, 1983). In a 1992 article, Slavin reports that children who work in groups have more positive feelings about themselves than do students in traditional classes.

2 Sharan (1990) suggests that by providing students with an active role in learning, group work increases students' motivation to learn. Co-operative learning has been shown to improve students' attitudes towards school and themselves as learners, as well as towards learning specific subject matter and learning in general (Johnson, Johnson & Scott, 1978, cited in Nastasi & Clements, 1991; Johnson, Johnson & Stanne, 1985; Slavin, 1980).

3 Group work also fulfils an important function in the socialisation of students (Sharan, 1990). It provides them with the opportunity to practise social processes such as the distribution of participation in a group context (Damon, 1984).

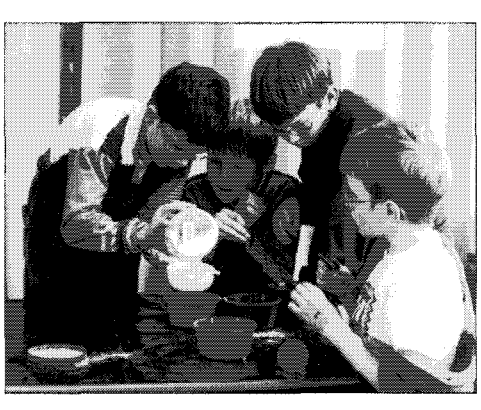
4 Working in groups fosters a pro-social orientation in students, which is manifested in an increase in concern for others, the enhancement of peer relations and a greater acceptance of minority and handicapped students (Bossert, 1989; Johnson & Johnson, 1985, cited in Nastasi & Clements, 1991; Slavin, 1980; Slavin, 1990a).

5 Involving students in equal status interaction leads to positive inter-group relations (Sharan, 1980; Slavin, 1987, 1990a, 1992). Moreover, when students share information and experience mutual assistance and joint pleasure in a common achievement, they can foster respect for

others with whom they might otherwise never interact (Edwards, 1994). Hence, group work can help students of different races, cultures and genders see one another as people rather than as members of distinct groups. As a consequence, small groups have been widely recommended as a means of achieving equity (see, for example, Oakes & Lipton, 1990).

6 Group work provides students with the opportunity to develop the interpersonal skills needed to work effectively with others. These skills are necessary in a range of situations in their student and adult lives. Specifically, group work helps students gain the experience of working in a setting where group members share common goals. It helps them to accept joint responsibility and to work with others to maximise the performance and output of the group (Webb, 1994; Yeomans, 1983).

7 Sharan (1976), cited in Cohen (1986), argues that having students make decisions on their own rather than telling them what to do has a desirable socialising effect on them, especially in political terms. Students not only will have a greater sense of control over their own environment but also will learn how to be active citizens in a collective rather than in an individualistic sense.



FACTORS THAT INFLUENCE GROUP WORK

A number of factors relating to the structure of the group have been identified as influential in group work. A range of these factors (except for group gender composition, which I discuss later) is considered below.

GROUP SIZE

Bossert, Barnett and Filby (1984) refer to a number of studies which show that instruction is more productive in small than in large groups. They report that students who receive instruction in small groups experience more individual assistance, more positive feedback and greater exposure to a wider variety of materials than do students who receive instruction in large groups. For example, in one of the studies that they cite (Peterson, 1981), students of both high and low ability participated more frequently in the activity in small group formats than did similar students in larger group settings. Students also retained more information when instruction occurred in small group formats.

Kutnick (1994) suggests that interaction involving all of the group members is more likely in small groups than in large groups. In a research review, Levine and Moreland (1990, 593) conclude that

> as a group grows larger, it also changes in other ways, generally for the worse. People who belong to larger groups are less satisfied ... participate less often and are less likely to co-operate with one another.'

In addition, in larger groups there is more chance of 'diffusion of responsibility' or 'social loafing (Webb, 1989)'.

This occurs when certain group members believe it is not their responsibility to initiate and carry out the activity and therefore sit back and let others do the work (Slavin, 1990b). Such an occurrence is detrimental

for group productivity in situations where the successful completion of tasks depends on the active contribution of the different group members.

THE NATURE OF THE TASKS

Cowie and Rudduck (1989, 162) categorise group tasks into three types:

- *Discussion tasks:* These may focus on the interpretation of something ambiguous, the sharing of experience, the pooling of ideas or the eliciting of opinions on an issue of common concern. They may require negotiation in the interest of arriving at a group consensus.
- *Problem-solving tasks:* These usually depend on the discussion of alternatives as a medium for constructive interaction.
- *Production tasks:* These are slightly different from problem-solving tasks in that there is usually a concrete outcome.

The three types of tasks do not necessarily involve the same group processes. Bennett (1994a) explains that although there is agreement that task characteristics are important and powerful mediators of group processes, little is, as yet, known of their effects. Studies usually provide insufficient detail of the tasks used and the demands made on the groups. Various conceptualisations have been suggested for the structure of tasks — from closed to open (Willems, 1981) and from tight to loose (Barnes & Todd, 1977). Bennett (1994b) observes that although the labels are different, the distinctions are similar in meaning, with the closed, tight end of the continuum defining tasks that are clearly specified, have one solution and require low-level thinking.

One study that comments on the structure of the task was carried out by Crozier and Keinberg (1987), quoted in Galton and Williamson (1992). This study, which involved 7- to 11-year-olds working on mathematics problem-solving tasks, reported that, contrary to what the researchers had expected, open-ended problems tended to be less helpful in stimulating collaborative discussions. In such cases, where children were expected to find an acceptable solution rather than the correct answer, the groups tended to agree on the first suggestion put forward by one of the members. Galton and Williamson concluded that 'problem-solving tasks with a clear testable outcome tended to generate a greater degree of collaboration than more open-ended tasks' (43).

PERSONAL CHARACTERISTICS OF THE GROUP MEMBERS

AGE

Kutnick (1988) points out that the development of co-operative skills takes place throughout the years of schooling. Older children therefore are expected to have better group skills than younger ones. However, he argues that children's ability to understand and interact with other children is not dependent on their age but on their experience and culture.

Yeomans (1983, 103) notes that 'very little documented research has examined the capacity of young children to explore and handle ideas and to generate knowledge through group discussion and interaction'. She reports Tough's (1977) work, which attempted to use collaborative groups for facilitating the language development of nursery and infant children. This work showed that the children were capable of engaging in extended exploratory discussion to some extent. Webb (1983) also has reported that children as young as seven could benefit from the experience of group work. I did not come across empirical stud-

ies that compared this experience at different ages or that examined the manner and extent to which it changed over time.

PERSONALITY

Very little research has been undertaken into the role of personality in the group work experience. Kutnick (1994) suggests that extrovert personalities are more likely to interact in small groups and introverts are less likely to interact. Indeed, Hare (1992) has related introversion-extroversion to the extent to which someone is likely to participate in group settings. Two of three studies that have included the personality of students in their investigation of individual characteristics, students' interaction and achievement in different types of groups report a relationship between personality and the group experience. These studies (Webb, 1982a, 1982b revealed that in group interactions, extroverted students tended to be ignored less frequently than introverted students when requesting help from other group members. In contrast, the third study (Webb, 1984) found no relationship between extroversion and introversion and student interaction in group situations.

ABILITY

Webb (1991) reviewed nine studies that correlated individual student ability with group interaction variables. While most of these studies used the students' scores on an ability test as the ability score in the analysis, some also examined the students' relative ability within the small group. Relative ability was typically defined as the difference between a student's score and the mean score of the group. Webb reported that most of these studies found that the high-ability students tended to give more explanations and information. However, the studies that examined relative ability within the group found that it was this ability, not absolute ability, which determined the extent of involvement of the different group members. This review also reported that the low-ability students were off-task more often than the high-ability students.

The few studies that have compared group processes in different group compositions show that students with different abilities perform better in particular types of groups. Two studies (Webb, 1982b; Webb & Kenderski, 1984) that examined groups with a wide range of abilities (high, medium, low) found that, in many of the groups, the high-ability and low-ability students formed a teacher-learner relationship while the medium-ability students tended to be left out of the group interaction, participating less than the highs and the lows. In contrast, the medium-ability students were very active in homogeneous medium groups. However, in mixed-ability groups with a narrower range of abilities (high-moderate, moderate-low) all students tended to be active participants. These findings were replicated by Bush (1997).

GROUP COMPOSITION

COMBINATIONS OF ABILITIES

A number of studies have investigated the relationship between the ability combinations of the group members and the group processes. Research by Webb (1989, 1991) and Bennett and Cass (1988) found that homogeneous groups of high-ability, medium-ability and low-ability students did not share an identical group experience. Webb found that only the homogeneous medium- ability groups showed high-level elaborative interactions that supported achievement during the task of jointly solving a mathematical problem. Students in high-ability groups did not display high-level elaborative interactions; most of them wanted to work as individuals. In a different subject area, Bennett and Cass found that the homogeneous high-ability groups significantly and consistently outperformed

the other types of homogeneous groups when performing a history task. Although these studies do not agree with regard to which group type provides the most productive experience, they do agree that the experience is not consistent across the different types of homogeneous groups.

Research relating to the experience of homogeneous low-ability groups has been consistently negative. Webb (1989, 1991) reports that these groups provide little stimulus (from more knowledgeable group members) for high-level elaborative interactions and that much of their interaction is off-task. Similarly, Good and Marshall (1984) found that these groups were prone to interruptions, spent less time on task and were very passive in the learning process. Furthermore, Bennett and Cass (1988) report that the experience is notably less productive for this group type.

Researchers have also compared the processes that occur in homogeneous groups, heterogeneous groups (with a combination of high-, medium- and low-ability students) and mixed-ability groups (with a combination of students from two ability levels). Webb (1984) reports that students in groups with a range of two abilities (high–medium or medium–low groups) tend to give more explanations than students in the other group types. Similarly, Bennett (1988) reports that mixed-ability groups interact more than heterogeneous groups and homogeneous groups. According to Bennett, the mixed-ability groups both gave and requested more explanations than the other group types and they also provided the most suggestions. The studies described here suggest that while mixed-ability groups (with two ability levels) are optimal for all students, heterogeneous groups (with three ability levels) benefit some students while homogeneous groups benefit others.

RACE AND ETHNICITY

Webb and Palinscar (1996, 859) argue that 'although group work may promote face-to-face contact among students from different groups, the condition of equal-status interaction may be difficult to fulfil'. One reason for this is that race and ethnicity may serve as diffuse status characteristics. According to expectations states theory, when group members do not know one another well, they can form judgements about each other using socially evaluated characteristics such as race and ethnicity (Berger, Rosenholt & Zelditch, 1980).

Studies that have investigated group interaction in multi-racial groups report that group members do not participate equally, especially in laboratory conditions where they do not know one another well. Webb (1982c) summarised the findings of seven of these studies as follows: white students tend to be more active and influential than minority students, while minority students tend to be less assertive and more anxious, talk less and give fewer suggestions and less information than white students. Cohen (1984) summarised the results of four studies in which students of different racial backgrounds worked on a collective task in four-person groups. She concluded that while playing a board game, whites were more dominant than blacks, Chicanos and American Indians, and that in Israel, Jews of Western origin were more dominant than Jews of Middle Eastern origin. In another study carried out in Israeli classrooms, Sharan and Shacher (1980), quoted in Cohen (1994), gave mixed racial groups of Middle Eastern and Western Eastern Jews a discussion task. The researchers observed that while performing the task, Western Jews took significantly more turns at speaking than the Middle Eastern Jews and used significantly more words per turn.

SUMMARY

The above discussion shows that the makeup of a collaborative group has profound implications for the experiences of the group members. It also shows that groups can vary on a number of variables simultaneously so that it is difficult to uncover the relative impact of each of them separately. Moreover, it also highlights the need for studies on group work to acknowledge the interaction of variables on the group experience and hence to focus on examining the relationships among these variables.

2

SETTING THE AGENDA FOR THE PROBE STUDY

In this chapter I give an overview of the National Education Monitoring Project and describe the position of the current probe study within it. I then explain the design of the probe study, present its research questions and describe the research methods used to answer them.

THE NATIONAL EDUCATION MONITORING PROJECT [1]

New Zealand's National Education Monitoring Project (NEMP) commenced in 1993 with the task of assessing and reporting on the achievement of New Zealand primary school children in all areas of the school curriculum. Children are assessed at two class levels: year 4 (half way through primary education) and year 8 (at the end of primary education). Different curriculum areas are assessed each year, over a four-yearcycle. (**Table 1** provides the schedule for the assessment of the different curriculum areas.) This four-yearcycle also incorporates the assessment of attitudes and the following skills—communication, problem-solving, self-management and competitive, social and co-operative, and work and study.

The NEMP provides a snapshot of what children can do at year 4 (ages 8-9) and year 8 (ages 12-13). The information is gathered from random samples of students. The national samples, containing 1,440 students at each of the two levels, represent about 3 percent of the children at those levels in New Zealand schools. The 1,440 students selected at each level are divided into three subgroups—A, B and C. These 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 stu-

TABLE 1 SCHEDULE FOR CYCLE 1 OF THE NATIONAL EDUCATION MONITORING PROJECT

YEAR		NEW ZEALAND'S CURRICULUM FRAMEWORK
1	1995	Science, art, information skills (graphs, tables and maps)
2	1996	Music, aspects of technology, language: reading & speaking
3	1997	Mathematics, social studies, information skills (library, research)
4	1998	Language: writing/listening/viewing; health & physical well-being

[1] The information in this section is adapted from Crooks and Flockton (1996).

dent 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 process and task achievement. Tasks requiring higher levels of professional judgement, such as group tasks, are marked by teachers. Tasks that can be marked objectively or with modest amounts of professional experience are marked by senior tertiary students. The results are analysed and reported task by task, with consideration being given to such variables as student gender and ethnicity, the geographical zone within which each school falls, and the socio-economic indexes of the schools. However, the emphasis of the reports is on the overall national picture of what New Zealand children can do at years 4 and 8.

THE CASE FOR GROUP ASSESSMENT

The NEMP is one of the few large assessment projects world-wide 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 collaboration into performance assessment settings. These include the need to achieve the following:

- 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 is presently unclear whether any type of group composition is 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) suggests that 'certain groups may be unfair if they do not give students equal access to favourable group processes' (249). She advocates the need for research that 'explore[s] 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'(255). My study responds to that need.

WHAT CAN BE ASSESSED IN GROUPS?

Group assessment may focus on (a) the processes that occur during the activity and/or (b) the group product/s. Either of these can be assessed in turn at the level of individual students and/or the whole group (Cowie & Rudduck, 1989). I discuss these options below.

PROCESSES

The assessment of processes focuses on the cognitive and/or social behaviours that students engage in during collaborative group work (Bennett & Dunne, 1992a). Although it is convenient to discuss cognitive and social processes separately, the distinction is by no means clear-cut, as Bennett and Dunne (1992a, 172) point out:

> the processes of group work tend to be so complex and overlapping that they cannot necessarily be separated out ... the relationships between the cognitive and social demands of a task may be so intricately interwoven that observation of separate features is not [always] possible.'

Because of the overlap between cognitive and social processes, it is necessary to describe what is generally understood to be included in the assessment of these processes.

COGNITIVE PROCESSES

The assessment of cognitive processes focuses on the attributes of thinking, reasoning, knowledge and understanding that group work promotes in students (Light & Littleton, 1994). Insight into these cognitive processes can be obtained through the study of students' verbal interactions. As Webb (1995, 2) argues,

> research on collaborative group work in the classroom shows that students verbalise their thinking in the process of helping one another, working together to solve a problem or complete a task, resolving disagreements, and justifying their actions, strategies and decisions.

Thus, when assessing a group's cognitive processes, account is taken of the following: the content of students' explanations, discussions and arguments in the course of completing a task; the manner in which they read, interpret and reword task instructions; and their evaluations of one another's ideas and their co-construction of ideas.

SOCIAL PROCESSES

Group assessment can also focus on the social processes that occur. These, as various researchers have noted (e.g., Bennett & Dunne, 1992a; Johnson & Johnson, 1979; Schmuck & Schmuck, 1983; Slavin, 1987; Thousand, Villa & Nevin, 1994; Webb, 1997), include the extent and nature of

- the individual group members' levels of involvement in task-related behaviour
- their ability to interact as a group by listening, responding to others and building on one another's ideas
- their willingness to elicit and provide help
- their co-operative behaviour in terms of sharing the task
- conflict and controversy.

Thus, the assessment of social processes can examine students' *participation* on the task, the extent to which such participation is *interactive* and, in turn, the extent to which interactive behaviour (verbal and non-verbal) during group work is *co-operative* and/or *conflict-oriented*.

PRODUCTS

Group work can also be assessed in terms of the product, that is, the outcome of the collaborative process (e.g, a completed worksheet, a model, a drawing). Not all group tasks result in an assessable product (e.g., discussion tasks where records of students' ideas and/or conclusions are not obtained). In such cases, the quality of the group effort must be assessed through an analysis of the kinds of processes described above.

INDIVIDUAL PRODUCTS

Students may each be assessed on the basis of a separate piece of work subsequent to the group activity. For example, after an experiment, each student completes an individual report or test and is assigned an individual mark on the basis of his or her performance on that task. This manner of assessment is recommended as a way of checking if individual students are making progress (Cohen, 1986). Although this approach is straightforward and makes individual students accountable for their work, some commentators have criticised it:

1 Bennett & Dunne (1992b) contend that it may encourage students to copy from or to depend heavily on other students' ideas.

2 Cowie & Rudduck (1989) and Webb (1997) state that it may work against the actual processes that group work is meant to promote. In other words, by encouraging students to think about their own individual assessment, this approach militates against the co-operative spirit that group work is meant to instil in students. Support for this assertion is also provided in research on the nature of rewards, which found that rewarding students individually discourages group work (Slavin, 1990a; Webb, 1989).

JOINT PRODUCTS

Students' collaborative work may also be assessed on the basis of a joint product. This joint product can be assessed either individually or jointly. The former is difficult to achieve because individual accountability in relation to joint products is low, and establishing the contribution of individual children is problematic (Bennett & Dunne, 1992a). Assessing the product jointly (i.e., assigning one mark or reward to the whole group) has been shown to be an effective way of promoting co-operative behaviour, for example, by promoting peer interaction and by encouraging students to help one another or elaborate on their contributions (Johnson, Johnson & Stanne, 1985; Sharan, 1980; Slavin, 1980, 1983). Despite these advantages of joint products, Slavin (1987, 1164) points out that when the whole group completes one product,

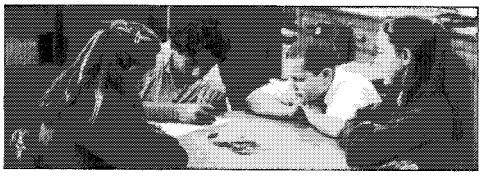
> there is a danger that some group members' efforts will not be needed or may even interfere with the group's success. For example, in a heterogeneous four-member group, the one or two most able students could probably complete a group worksheet by themselves as well or better than if they actively involved the less able group members.

In response to this observation, however, it may also be said that such occurrences are less likely when collaborative work engages students in true group tasks that are ill-structured. According to Cohen & Arechevala, cited in Cohen, 1994, 8), ill-structured tasks are those that require

> resources (information, knowledge, heuristic problem-solving strategies, materials and skills) that no single individual possesses so that no single individual is likely to solve the problem or accomplish the task objectives without at least some input from others.

SUMMARY

In summary, the literature on group assessment indicates that if we are to gain a comprehensive understanding of such work, we require insight into not only its end-products but also the inter-related processes that lead to these products. In addition, we need to focus on the groups' collaborative products and not solely on the individual contributions of group members.



THE PROBE STUDY

Given that the 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 the 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, 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 girls with a similarly enjoyable and productive experience while undertaking three tasks from different areas of the New Zealand curriculum. 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 composition, the two age levels of the students, and the different nature of the three tasks.

THE GROUP TASKS

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

1 A 1995 science task called *Separating Mixtures*, which is made up of the following steps:

The children first discuss how to separate a mixture made up of four substances (flour, paperclips, polystyrene balls and ball bearings). They come

up with a plan and share it with the teacher, who writes it down for them.

They are then given the following equipment (sieve, magnet, funnel, tweezers, magnifying glass) and asked to think about their plan in the light of the equipment available. They have another discussion in order to see what modifications they need to make to their plan, and they then share this plan with the teacher.

Once the teacher has written down the modified plan, the students are asked to go ahead and separate the mixture.

After the experiment, the students are asked to evaluate their plan and to explain what modifications they would make to the experiment if they had the opportunity to do it again.

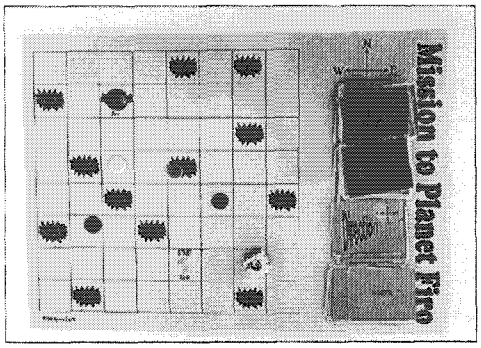
2 A 1996 language task called *Question Time*, made up of the following steps:

The students are told to prepare for a visit from two police officers (one male and one female) by thinking up questions they could ask them about their work and their lives.

The students have a brainstorming session where they think up questions and communicate them directly to the

teacher, who writes them down on a sheet of paper.

- With their list of questions available, the students are asked to work as a group to determine the six questions they will ask the police officers. The students are then left on their own to get on with the process of choosing the questions.
- When they are ready, the students inform the teacher, who asks them to report back on their chosen questions and to give reasons for their choices.



3 A 1996 technology task called *Space Game,* made up of the following steps:

- The students are told that they will be playing a board game and that during it they are (a) to think about ways of improving it, and (b) to think up questions to ask other people to find out what they think about it.
- They watch a video, which gives them instructions on how to play the game. They are then asked to play the game for five minutes and to think about ways of making it better. They are told that the game needs to be fun for children who are six years or older and that through the game the children need to learn the directions north, south, east and west.
- After playing the game, the children have a five-minute discussion on how to make the game more fun. At the end of this discussion, they choose their four best ideas and report them to the teacher.
- They then have another five-minute discussion on how they could find out if other children like the game and whether other children think it needs improving. At the end of this discussion, they again choose their four best ideas and report them to the teacher.

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 Do the children's evaluations of their NEMP group experience vary across the different 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. **Table 2** presents the schedule for gathering the raw data, and **Figure 1** illustrates the relationship between the student samples and the different research methods used.

FIGURE1 RESEARCH TOOLS AND STUDENT SAMPLES

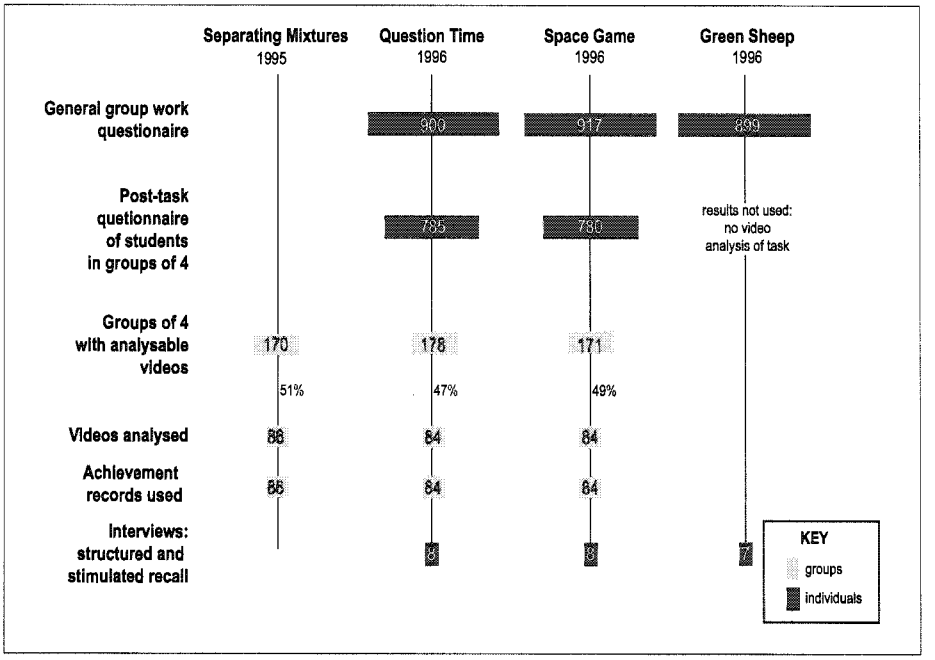


TABLE 2 SCHEDULE FOR GATHERING THE RAW DATA

RESEARCH TOOL	DATA COLLECTION PERIOD
Videotaped recordings	July 1995 – October 1996
Questionnaire	August 1996 – October 1996
Interviews	September 1996 – October 1996
Achievement records	January 1996 – February 1997

VIDEOTAPED RECORDINGS

The videotaped recordings of the four-member groups of children working on the NEMP collaborative tasks gave me the opportunity to analyse the group processes of a large number of groups and the possibility to review the processes in any one group as many times as I needed.

During the NEMP group tasks, the children generally work without direct supervision and tight control from the teacher administrators, as is normally the case in natural classrooms. Therefore, reactivity (i.e., the effect of the administrators on the students) is expected to be relatively low. However, the children know that they are being video-recorded and that they are not in their usual learning environment. While these conditions are common in a performance assessment setting, some caution is needed when applying the results of this study to the classroom setting.

The choice of techniques available for analysing group processes includes categories, checklists, rating scales and narrative accounts. I did some trial work using categories and checklists, but found both techniques unsatisfactory. On the one hand, categories gave a microscopic insight into a group's processes in a way that made it difficult to get an overview of what had happened in the group. On the other hand, checklists were not accurate enough, as they only enabled me to record whether a particular behaviour was present or absent at particular points in time. Consequently, I developed a structured observation schedule using rating scales and narrative accounts. I now explain each of these in turn.

RATING SCALES

The components of rating scales are (a) the listing of the dimensions to be rated, and (b) the scale (referred to as a Likert scale) for rating each dimension (Gredler, 1999). For example, in this study, participation on a task was a dimension, and the scale for that dimension was (1) never, (2) rarely, (3) half of the time, (4) most of the time, (5) always. Co-operation was another dimension, the scale for which was (1) low, (2) moderate, (3) high.

The use of rating scales as a coding strategy allowed me to incorporate both the frequency and intensity of the aspects of behaviour I was interested in observing (participation, interaction, co-operation and conflict). Fassnacht (1979, 136) explains this benefit as follows: 'ratings are used to conceive of behaviour as a whole, dimensionally or to quantify it in terms of intensity and frequency.' I also considered rating scales appealing because they enabled me to make a decision on the behaviours after periods of viewing rather than throughout the process. As Rosenshine and Furst (1973, 132) note, the use of rating scales allows 'the observer ... to estimate the frequency of specified events or constellations of events only once, usually at the end of an observation session'. Therefore, by using rating scales, I was able to process many cues before making decisions. I soon became aware, however, of the need to define terms, such as participation, clearly so that the scales would be representative of the behaviour under study. I also needed to be able to differentiate among the different levels of the ratings scales. As Stallings and Mohlman (1988, 471) explain, 'in order to produce usable data, very specific definitions must be made of the attributes at each point of the scales'. I did this by viewing the videotapes several times, identifying video exemplars and/or producing and modifying descriptors for the different levels of each rating scale and then checking all the tapes against these video exemplars and/or descriptors.

NARRATIVE ACCOUNTS

Because of the large number of tapes I was interested in analysing, it was not feasible to transcribe all the tapes and then make a retrospective analysis of specific aspects of the total recording. Consequently, I decided on the use of narrative accounts. This strategy allowed me to write a running commentary of what was happening in each group for the different activities that made up each task, for example, during the discussion and then during the experiment of the science task *Separating Mixtures*. Stallings and Mohlman (1988, 473) explain the main advantage of recording data in this way:

> the context can be described in a rich and holistic manner. The natural sequence of events is preserved. Unpredicted events can be recorded. Qualitative statements can be made ... none of the quantitative observation instruments could adequately record that kind of information.

This tool helped me to remain focused while viewing the tapes, and it provided me with another source of data when I needed to decide on the levels of the rating scales for the different processes. As tools, the narrative accounts and the rating scales complemented each other because the former needed to be recorded while I was viewing the tapes whereas the latter needed to be completed after the viewing.

THE VIDEO ANALYSIS PROCESS

I originally planned to have, for each of the three tasks, 10 groups for each of the five group types (4b, 3b1g, 2b2g, 1b3g and 4g) at both years 4 and 8. However, as a result of the random sampling technique used in the NEMP, it was rare to have 10 same-gender groups of boys and of girls. Therefore, when the number of groups available was

TABLE 3 PROPORTION OF TAPES ANALYSED FOR THE THREE TASKS

	TASK	PROPORTION	%
Science	*Separating Mixtures*	86/170	51
Language	*Question Time*	84/178	47
Technology	*Space Game*	84/171	49

less than 10, I included all the tapes with four-member groups. When there were more than 10 groups, I used systematic sampling to make my selection. In all three tasks there were a number of tapes that could not be used, mainly because group members were absent or there were recording problems. **Table 3** shows the proportion of tapes that I analysed in relation to the number of complete tapes available for the analysis of each task. In **Table 4** and **Figure 2** the data are broken down by group type and age level.

I initially developed a structured observation schedule that captured a number of the processes present in the first task (*Separating Mixtures*). I then modified the schedule for each of the other two tasks (*Question Time* and *Space Game*) in order to capture both the general and specific characteristics of the different tasks. The versions of the observation schedules used in the analysis were developed after trialing and modifying earlier versions. A computerised database for each of the three observation schedules was created using

FIGURE 2 PROPORTION OF TAPES ANALYSED BY GROUP TYPE AND AGE LEVEL

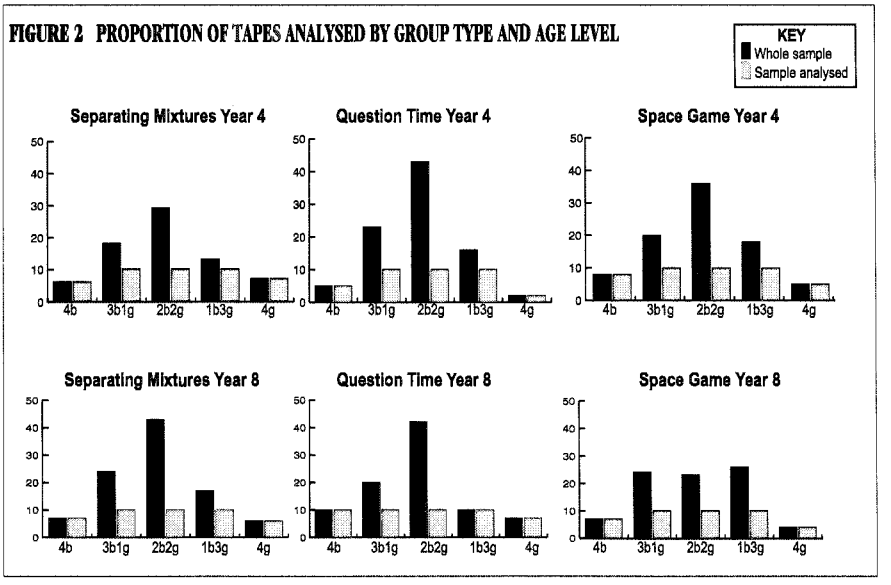


TABLE 4 PROPORTION OF TAPES ANALYSED BY GROUP TYPE AND AGE LEVEL

GROUP TYPES	SEPARATING MIXTURES				QUESTION TIME				SPACE GAME			
	YEAR 4	%	YEAR 8	%	YEAR 4	%	YEAR 8	%	YEAR 4	%	YEAR 8	%
4B	6/6	100	7/7	100	5/5	100	10/10	100	8/8	0	7/7	100
3B1G	10/18	56	10/24	42	10/23	44	10/20	50	10/20	50	10/24	42
2B2G	10/29	35	10/43	23	10/43	23	10/42	24	10/36	28	10/23	44
1B3G	10/13	77	10/17	59	10/16	63	10/10	100	10/18	56	10/26	39
4G	7/7	100	6/6	100	2/2	100	7/7	100	5/5	100	4/4	100
TOTAL	**43/73**	**59**	**43/97**	**44**	**37/89**	**42**	**47/89**	**53**	**43/87**	**49**	**41/84**	**49**

Filemaker Pro® (Claris, 1995) and CVideo® (Envisionology, 1993), with the help of Miriam Richardson, the production editor at the Educational Assessment Research Unit (EARU), University of Otago. This set-up made the video analysis process very efficient.

TABLE 5 CONDUCTING THE VIDEO ANALYSIS — HOURS SPENT ON EACH TASK

TASK	NUMBER OF VIDEOS	VIEWING TIME/ TAPE† (HOURS)	TOTAL VIEWING TIME (HOURS)
Separating Mixtures	86	4.00	344
Question Time	84	3.00	252
Space Game	84	6.00	504
TOTAL	**252**	**4.33**	**1,100**

†Average time for viewing each tape four times.

I analysed the videotapes in the order of the NEMP classification numbers rather than according to their group gender composition. This allowed me to reduce the possibility of developing bias in relation to particular group types. I found that it was not necessary to view a tape before coding it, so I coded it on the first viewing. For each task, I chose video exemplars and developed descriptors for the levels of the ratings 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 4 tapes and once starting with the year 8 tapes to ensure that I was as consistent as possible in similarly coding the tapes of the two age levels. **Table 5** shows the duration of the video analysis period broken down by task. It does not include the periods when I was developing and trialing the observation schedules.

THE CROSS-CODING

About 10 percent of the tapes (10 for each of the three tasks) were recoded by another researcher. Denzin (1970) refers to the process of more than one researcher studying the same phenomenon as investigator triangulation. The aim of this process was to explore the extent of agreement between two persons in interpreting the same events using the definitions of the variables and of the levels of the rating scales as well as the video exemplars that I had used.

I approached Robyn Caygill to assist me in this process. Robyn had worked as an educational researcher at the Ministry of Education and at the Educational Assessment Research Unit (EARU) and was familiar with both the content of the tasks and the process of administering them. For each task, the training consisted of three steps. First, we worked through the meaning of the particular categories and the levels of the rating scales, using notes and video exemplars. Second, Robyn coded two tapes in my absence. Finally, we checked our independent coding and made necessary clarifications. For each task, this process was done over two consecutive days (with one exception, when a weekend fell between the first and second days of training).

The tapes that were to be recoded were not chosen randomly. The selection deliberately included a range of levels for the major categories that I had coded. However, this information on the process of choice was not communicated to Robyn. The recoding of 30 tapes was completed over a period of nine weeks. It took Robyn 70 hours to view the tapes once and to code them accordingly.

Once I had copies of the group records that Robyn had coded, I identified the major categories and checked the extent of our agreement for coding the three tasks. Agreement was coded with (A), and disagreement was coded with (D). However, because a large number of rating scales had five levels, I also noted the level of disagreement. If Robyn had coded one level higher than I had, this was coded as (D+1).

If she had coded one level lower, this was coded as (D-1). The same strategy was used for disagreements at two levels or more. On average, we had 70 percent agreement on the coding. **Table 6** presents the agreement percentages of the cross-coding of the three tasks.

TABLE 6 EXTENT OF AGREEMENT BY RESEARCHERS IN THEIR CROSS-CODING OF THE THREE TASKS

	SEPARATING MIXTURES	*QUESTION TIME*	*SPACE GAME*	OVERALL
	%	%	%	%
Interaction levels	90.00	100.00	75.00	88.33
Co-operation levels	80.00	70.00	50.00	66.67
Conflict levels	40.00	97.00	73.00	70.00
Help levels	†	†	60.00	60.00
Interference levels	†	†	50.00	50.00
Mean group levels	**70.00**	**89.00**	**61.60**	**73.53**
Extent of all group ...				
participation	60.00	90.00	75.00	75.00
interaction	50.00	50.00	75.00	58.33
co-operation	65.00	60.00	45.00	56.67
Mean group involvement	**58.33**	**66.67**	**65.00**	**63.33**
Participation levels	83.75	89.00	62.75	78.50
Idea levels	62.50	100.00	74.00	78.83
Organisation levels	47.50	67.50	61.00	58.67
Mean individual levels	**64.58**	**85.50**	**65.92**	**72.00**
TOTAL MEAN	**64.30**	**80.39**	**64.17**	**69.62**

† Not coded in these tasks.

As is evident from **Table 6**, the range of agreement varied considerably for the ratings of the different scales for any one task as well as for the rating of the same scale for the different tasks. It is worth nothing that for each section, the level of agreement on one or two rating scales was higher than on others. This suggests that during the viewing Robyn tended to focus on certain group processes and, in so doing, did not observe the others as well as she might have otherwise. This is a common occurrence when the tapes are viewed once only and the different processes being examined are complex. It is for this reason that I present in **Table 7** the levels of disagreement.

ing, considering that she coded the different categories using one viewing whereas I viewed each tape four times. The cross-coding experience suggests that, in reality, it is impossible to focus on numerous processes accurately during one viewing and that, therefore, in a similar situation, it would be better to focus on a limited number of processes at any one time.

In **Table 7**, it is worth noting that for 84 percent of the cases of disagreement, the difference was for one level. For all three tasks, the majority of these cases involved Robyn coding at a lower level than I did. This was not surpris-

TABLE 7 EXTENT OF DISAGREEMENT BY RESEARCHERS IN THEIR CROSS-CODING OF THE THREE TASKS

	SEPARATING MIXTURES %		*QUESTION TIME* %		*SPACE GAME* %		OVERALL %
	–	+	–	+	–	+	
Level 1	57.00	23.00	69.00	19.00	44.00	41.00	84.33
Level 2	9.00	6.00	6.00	•	11.00	4.00	12.00
Level 3	1.00	4.00	6.00	•	•	•	3.66
Mean group levels	**67.00**	**33.00**	**81.00**	**19.00**	**55.00**	**45.00**	**99.99**

THE QUESTIONNAIRE

In order to explore how children feel about working in different group types generally and how they felt during one NEMP group experience in particular, I used a questionnaire with all the students who participated in the NEMP in 1996. The questionnaire served as a self-report for the children. Given that the questionnaire needed to be completed by both year 4 and year 8 students, the questions generally required a response to be circled rather than a written answer. Furthermore, in order to reduce the necessity for literacy skills, I used, where possible, a non-verbal answering scheme that used faces and stars. Writing was limited to writing the names of the members of the group, and this was required in two responses only.

The questionnaire was first trialed with two groups of year 4 and year 8 students who were not involved in the NEMP. It was then completed by all year 4 and year 8 children who participated in NEMP in 1996. The students who were assigned to Groups A completed the questionnaire after *Space Game*. Those in Groups B completed the questionnaire after *Question Time*, and those in Groups C completed it after *Green Sheep*. (This last task was planned for video analysis, but later was not included because of the large amount of data generated by the other two tasks and the *Separating Mixtures* task.)

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 students. The students then marked their own answers on their sheets. Help was at hand for any students who required further assistance in understanding the questions and/or in answering them. The questionnaires were completed in two five-week periods between August and October 1996 when the NEMP teacher-administrators were in the schools carrying out the assessments. There was a 100 percent return of the questionnaires that had been completed (i.e., 94 percent of the national sample participating in the NEMP in 1996).

Table 8 shows the number of students who completed the questionnaire at the two age levels. I include the responses of all these students when I report on how the children felt about working in groups with different gender compositions. I present the responses of the year 4 and year 8 students separately in order to see if there were similar patterns in the responses at the two age levels.

TABLE 8 NUMBER OF PARTICIPANTS WHO COMPLETED THE QUESTIONNAIRE

SAMPLE	BOYS	GIRLS	TOTAL
Year 4	720	650	1,370
Year 8	718	628	1,346
TOTAL	**1438**	**1278**	**2,716**

TABLE 9 NUMBER OF PARTICIPANTS INCLUDED IN THE ANALYSIS OF THE GROUP EXPERIENCE

SAMPLE		BOYS	GIRLS	TOTAL
Year 4	Groups A	219	189	408
	Groups B	208	192	40
	Subtotal	**427**	**381**	**808**
Year 8	Groups A	200	172	372
	Groups B	192	163	355
	Subtotal	**392**	**335**	**727**
TOTAL		**819**	**716**	**1,535**

Table 9 shows the numbers of 1996 year 4 and year 8 students in Groups A and B whose responses I included when reporting on the students' group experiences during a particular NEMP task. None of the questionnaire responses corresponded to *Separating Mixtures* because that was

a 1995 task (and I started this project in 1996). I kept the responses of Groups A and B separate, since the groups worked on different tasks. Doing so allowed me to see whether there were similar and/or different relationships between a particular task and the responses for the boys and/or the girls overall and/or for particular group types.

In order to analyse the questionnaire responses using a statistical programme, it was first necessary to translate the students' verbal responses into numbers. Student markers employed by the NEMP translated the students' responses into numbers using the coding schedule that I developed for this process. The data were entered into computer files by Computing Services, University of Otago. On receiving these files, I checked the data entries against the original questionnaire response sheets and filled in gaps where necessary.

I analysed the children's questionnaire responses using SYSTAT (Wilkinson, 1989). Frequencies and percentages were calculated for all the responses by age group and gender as well as by group gender composition. ANOVA one-way analysis of variance was used in the analysis.

THE INTERVIEWS

I was aware that the questionnaire constrained the expression of students' opinions and did not allow for unexpected or unanticipated answers. Also, it did not allow certain themes to be probed or developed in more depth (Burns, 1994). Because of these limitations, I interviewed a subsample of the students to explore their perspectives on group work.

My schedule allowed me to interview the 23 eight-year-olds who were randomly selected to participate in the NEMP year 4 assessments in the Dunedin city area in 1996. In one school (School 1), the children came from two classes, and in the other school (School 2), where there was one absent student, the children all came from the same class.

I interviewed the students in a one-off situation on the school premises. The time at my disposal meant that I had to interview four students in the course of one morning at times when they were not involved in other activities relating to the NEMP. I had a maximum of 45 minutes that I could spend with each student. In effect, the interviews lasted between 28 and 45 minutes, with an average time of 37 minutes. The differences in the duration of the interviews were the result of the differences in the students' elaboration of their responses.

Because of the time restrictions, I found it necessary to consider the option of group interviews. Although these would have had a number of advantages, I decided that they were not appropriate for my study because of two disadvantages highlighted by Fontana and Frey (1994, 374). The first concerned the risks of 'group think', that is, the danger of all members of the group adopting the same idea. The second concerned the possibility of the emerging group culture interfering with individual expression, where one child finds it hard to disagree with other opinions or to voice his or her opinion on something for fear of being seen as different. Because I was interested in the ways individual children feel in specific group contexts, and because it was possible that they would be reluctant to express themselves freely in front of other children, I opted for individual interviews.

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 the one-off interview situation. In this report,

however, I present only those responses from the segment of the structured interview that related to the final research question relevant to this report.

Writers like Hitchcock and Hughes (1993) express doubts about the usefulness of the structured interview with children because they feel it is unlikely that it will reveal the complex factors that shape their social worlds. However, in my experience, this interview technique does reveal information on how children feel about working in groups with different gender compositions. I found it helpful to assure the students that there were no right or wrong answers to what I was asking. And although all the questions were pre-planned, at the end of each section I included an open-ended question for anything else that students wanted to add. So, for example, after I had asked the pre-planned questions on how they felt about working in same-gender and mixed-gender groups, I said, *'Nick, is there anything else you would like to tell me about this ... ?'* This open-ended prompt allowed the children to explain or elaborate on issues of interest to them.

With the students' permission, all of the interviews were video-recorded. None of the students objected to being videotaped, and the casual way in which they sat and their body language during the interview suggested to me that they did not mind being video-recorded. The main advantage of having the interviews video-recorded was that, when I was transcribing the interviews, I had access to the students' facial expressions and body lang uage as well as their verbal responses. This was especially helpful when the children made facial expressions and/or gestures instead of giving a verbal response.

I analysed the interview data using NUD*IST® (Qualitative Solutions and Research Pty, 1995). In order to use this programme, I first studied the interviews and set up the coding categories. Then, using the programme, I went through the data and sorted them into the right categories. NUD*IST® was useful for both types of interviews. For the structured interview, in particular, this programme made it possible to list the 23 responses to each of the questions in order, with each interviewee's personal data and the relevant transcript line numbers appearing automatically with each response. It also facilitated the listing of responses by gender groups and by school when comparisons were necessary.

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 the NEMP and used by the teacher-markers who were involved in marking the NEMP tasks. In this analysis, I examined the patterns of achievement of the different group types, and through the use of the Pearson correlation coefficient explored the relationships between the group product scores and the group means for the various processes analysed.

SUMMARY

In this chapter I described the position of the current probe study within the larger NEMP project. The probe study was designed to explore the effects of group gender composition on the processes and products of group work at two age levels and on tasks from three subject areas. Having described the design of the study, its research questions and the methods employed to answer those questions in this chapter, I present the results of the study in the next four chapters.

3

PARTICIPATION IN THE FIVE GROUP TYPES

When students work collaboratively on a task, a number of different social processes may occur. As I pointed out in the previous chapter, these social processes may be studied in terms of students' *participation* in the task, the extent to which such participation is *interactive* and, in turn, the extent to which interactive behaviour (verbal and non-verbal) during group work is *co-operative* and/or *conflict-oriented.* In this chapter I report on the participation of the members in the different group types, and in the following chapter I focus on the three group processes under study (interaction, co-operation and conflict).

Effective group work calls for students' active involvement in task-related activity. While it may be possible for students to learn simply by observing others without interacting with them (Bandura, 1986; Schunk, 1987), most theories about learning in group contexts argue that students' active involvement is essential for the development of new knowledge, skills and understanding (Webb, 1994). Webb (1989) and Webb and Kenderski (1985), for example, have shown that students who learn best from group work are those who are actively involved with the cognitive content of the task. They also reveal that 'social loafers' (i.e., students who do not participate in group tasks but allow others to do the work) achieve less on related achievement tests. Other empirical evidence shows that students learn less from listening to or watching problem-solving than from active engagement in the same activities (Johnson, Johnson, Roy & Zaidman, 1985, and Weinstein & Bearison, 1985, both cited in Nastasi & Clements, 1991).

Although, within the context of the present study, participation refers to individual on-task involvement, I need to explain, before presenting the results, what participation specifically encompassed in the three tasks. Having done this, I present the results for the participation analyses in two ways for the five group types. I first present the results at the group level and then present the results for boys and girls as separate subgroups within the different group types. This manner of reporting makes it possible to evaluate whether the mean participation level was consistently higher or lower in particular group types. It also makes it possible to evaluate whether boys and girls as subgroups consistently participated more or less in particular group types or whether their mean participation levels were similar in the different group types.

PARTICIPATION EXPLAINED

CODING THE PARTICIPATION LEVELS

SEPARATING MIXTURES

The science task *Separating Mixtures* was made up of two activities—the discussion and the experiment. So that the task would not be fragmented into too many parts, the group dynamics of both discussions (without and with the equipment) were coded as one activity in the discussion section of the observation schedule. The individual students' participation did not include those ideas that were not shared in the group but were then reported directly to the teacher. Nor did it include reporting the plan to the teacher, because in many groups it was the teacher who decided on who reported back.

The coding of the experiment focused on the organisation and performance of the experiment, that is, until the time that the group finished separating the mixture or were told to stop. The cleaning-up was not coded as part of the experiment because it did not occur in all of the groups, and because in those groups where it did occur, it could have been the teacher who delegated the work. The students' evaluation of the experiment and the modifications that would be necessary next time round similarly was not coded as part of individual students' participation, because in many groups it was not the students themselves who decided on who did the reporting.

For both activities, individual participation was coded as 'absent', 'low', 'moderate' or 'high'. In the discussion, the code 'low' was assigned to students whose input was limited to one idea and/or very little other input. The code 'moderate' was assigned to students who contributed a few ideas and/or helped build the plan. The code 'high' was assigned to students who contributed highly either in terms of sharing ideas or putting the plan together or a combination of both. In the experiment, there were four jobs to be done. Therefore, the code 'low' was assigned to students whose input was marginal and involved less than one complete job. The code 'moderate' was assigned to students who either completed one whole job or else shared a number of jobs with other group members. The code 'high' was assigned to students who carried out two complete jobs or more.

QUESTION TIME

The language task *Question Time* was made up of three activities—brainstorming, question choice, and reporting and justification.

In the brainstorming activity, the individuals' participation level was assigned quantitatively without evaluating the content or the structure of the questions. A student who contributed one or two questions was assigned a 'low' participation level. A student who contributed three or four questions was assigned a 'moderate' participation level. A student who contributed five questions or more was assigned a 'high' participation level.

In the question choice activity, the quality of the contributions was taken into account in assigning the level of participation (e.g., well-articulated contributions counted more than nods). Video exemplars were selected for each level of participation. The code 'low' was assigned to students who chose one question by themselves, or partly contributed to the discussion of up to two questions, or contributed only through reading the questions aloud or ticking the group's choices. The code 'moderate' was assigned to students who chose two or three of the six questions by themselves, or were involved in the discussion of two chosen questions and ticked

the choices, or partly contributed to the discussion of three to four chosen questions (whether or not they were also involved in ticking the choices). The code 'high' was given to students who made considerable contributions to the discussion of at least five chosen questions (with or without contributing to reading or ticking the choices) or to students who chose four or more questions by themselves.

In the reporting and justification activity, the individual students' overall participation level in both reporting and justifying the choices was evaluated. The total of six questions to be reported and justified meant that students could make up to 12 contributions in this activity. The code 'low' was assigned to a student who made one or two contributions, the code 'moderate' to those who made between three and eight contributions, and the code 'high' to those who made nine or more contributions. A student who was highly involved in justifying all six questions was also assigned a 'high', irrespective of whether there was input from other group members.

SPACE GAME

The technology task *Space Game* was made up of a board game and two discussions. I present the results for the discussion phases of the two discussion activities separately because the content of each activity was different (one was on how to improve the game and the other was on how to market the game and find out from others if it needed improvement). In both discussions, participation involved the sharing of ideas, and acknowledging, questioning and building on the ideas offered by others.

The code 'low' was assigned to students whose overall input was limited to sharing or discussing up to two ideas. The code 'moderate' was assigned to students who shared, added to and/or discussed between three and five ideas. The code 'high' was given to students who contributed significantly to the discussion of six or more ideas. If a student's contribution was limited to nods and/or yes and no, the level assigned was one lower than it would have been if these contributions had been more elaborate.

CONDUCTING THE ANALYSIS

This analysis of participation was carried out using the following procedure. Having identified the individual students' participation level as absent, low, moderate or high, I translated these ratings into numbers using the following key: absent = 0; low = 1; moderate = 2; high = 3. For the overall mean participation levels in the different group types, I used this formula: group aggregate/4 (group members)/ number of groups. And for the separate mean participation levels for boys and girls, I modified the formula to correspond to the numbers of boys and girls in the particular group types: group aggregate for that gender/group members of that gender/ number of groups. The means were always out of a maximum of three.

RESULTS

1 MEAN PARTICIPATION LEVELS IN THE DIFFERENT GROUP TYPES

The participation means of the different group types for the various activities that made up the three tasks are presented in **Tables 10 and 11** for Years 4 and 8 respectively.

As the tables show, there was little discernible difference in participation between the same-gender and the mixed-gender groups at either Year 4 or Year 8. The level of participation was not consistently higher in the same-gender groups and in the 2b2g groups than in the same-gender groups.

TABLE10 YEAR 4: MEAN PARTICIPATION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	1.88	1.60	1.58	1.65	1.46
	Experiment	2.08	1.88	1.95	1.95	2.04
Language	Brainstorming	1.90	2.03	2.03	2.00	1.38
	Choice	2.25	2.23	2.25	2.28	1.88
Technology	Discussion one	1.88	1.60	2.13	1.93	1.85
	Discussion two	1.17	1.28	1.30	0.98	1.40
AVERAGE		**1.86**	**1.77**	**1.87**	**1.80**	**1.67**

TABLE11 YEAR 8: MEAN PARTICIPATION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	1.71	1.60	1.55	1.68	1.92
	Experiment	2.00	1.98	1.83	2.05	2.21
Language	Brainstorming	2.10	2.03	1.93	2.15	1.96
	Choice	2.55	2.30	2.18	2.45	2.36
Technology	Discussion one	1.82	2.50	2.15	2.70	2.88
	Discussion two	1.11	1.45	1.13	1.48	1.92
AVERAGE		**1.88**	**1.98**	**1.80**	**2.09**	**2.21**

Moreover, participation in the three mixed-gender group types was neither consistently higher nor lower than in the same-gender groups. It is worth noting that age influenced particular group types differently, with the members of the 4g groups participating more in the older age group than in the younger age group and the members of the 2b2g groups exhibiting the opposite pattern.

2 MEAN PARTICIPATION LEVELS OF BOYS AND GIRLS IN THE DIFFERENT GROUP TYPES

The analysis of participation also examined the mean participation levels of boys and girls in the different group types. This analysis related to the first research question of this study, which asked whether the task involvement of boys and girls would differ according to the gender composition of the group and whether it would change according to the children's age level and the nature of the tasks. The results of this analysis across the three tasks are presented in **Tables 12 and 13** for Years 4 and 8 respectively.

The mean participation levels for boys and girls as separate subgroups indicated a relationship between the gender composition of the group and the involvement of boys and girls in the tasks. This relationship changed with age. At Year 4, the mean participation levels for boys and girls were lower in the group type where they were outnumbered than in the other group types. Boys in the 1b3g groups participated less than boys in the other group types, while girls in the 3b1g groups and in the 4g groups participated less than girls in the 2b2g and the 1b3g groups. At Year 8, both gender groups, the girls especially, participated less in the 2b2g groups than in the other group types.

At Year 4, girls participated more in the mixed-gender group types where they were not outnumbered than in the same-gender groups. Conversely, at Year 8, girls' participation was greatest in the same-gender groups. The overall results of this analysis at Years 4 and 8 suggested that the disparity between the participation means of boys and girls was smaller in the gender-balanced groups than in the gender-imbalanced groups. However, there was no evidence of domination by males in the mixed-gender groups across the three tasks. Rather, the overall results at both ages suggested that the girls generally participated more than the boys, although the differences were small at Year 8. At both ages, there was no indication that girls were disadvantaged compared to the boys in the mixed-gender groups.

TABLE 12 YEAR 4: MEAN PARTICIPATION LEVELS BY GENDER AND GROUP TYPE

		4B	3B1G		2B2G		1B3G		4G
		B	B	G	B	G	B	G	G
Science	Discussion	1.88	1.67	1.40	1.50	1.65	1.50	1.70	1.46
	Experiment	2.08	1.97	1.60	1.70	2.20	1.60	2.07	2.04
Language	Brainstorm	1.90	2.00	2.10	2.25	1.80	1.80	2.07	1.38
	Choice	2.25	2.13	2.50	2.05	2.45	1.90	2.40	1.88
Technology	Discussion one	1.88	1.77	1.10	2.30	1.95	2.10	1.87	1.85
	Discussion two	1.17	1.33	1.10	1.30	1.30	0.90	1.00	1.40
AVERAGE		**1.86**	**1.81**	**1.63**	**1.85**	**1.89**	**1.63**	**1.85**	**1.67**

TABLE 13 YEAR 8: MEAN PARTICIPATION LEVELS BY GENDER AND GROUP TYPE

		4B	3B1G		2B2G		1B3G		4G
		B	B	G	B	G	B	G	G
Science	Discussion	1.71	1.63	1.70	1.50	1.60	1.40	1.77	1.92
	Experiment	2.00	2.03	1.90	1.80	1.85	2.10	2.03	2.21
Language	Brainstorm	2.10	1.90	2.40	1.90	1.95	1.80	2.27	1.96
	Choice	2.55	2.27	2.40	2.05	2.30	2.20	2.53	2.36
Technology	Discussion one	1.82	2.53	2.40	2.25	2.05	2.40	2.80	2.88
	Discussion two	1.11	1.43	1.50	1.15	1.10	1.50	1.47	1.92
AVERAGE		**1.88**	**1.97**	**2.05**	**1.78**	**1.81**	**1.90**	**2.15**	**12.21**

3 MEAN IDEA LEVELS GENERATED BY THE DIFFERENT GROUP TYPES

As well as looking at general participation, I also looked at the levels of ideas shared in the different group types during those activities involving the generation of ideas. I carried out this analysis in order to investigate whether certain group types generated higher or lower levels of ideas on average.

SEPARATING MIXTURES

Students who did not contribute any ideas during either of the discussions in *Separating Mixtures* (i.e., in the discussions without and then with the equipment) were assigned the code 'absent'. Students who contributed mainly one idea in either of the discussions were assigned the code 'low'. The code 'moderate' was assigned to students who contributed several ideas in either or both of the discussions (two to four ideas overall). The code 'high' was assigned to students who contributed ideas throughout the discussion (five ideas or more). Although the number of ideas was used as a guide, the quality of the ideas was also taken into account in assigning the levels.

QUESTION TIME

Students who did not contribute any questions during the brainstorming activity of *Question Time* were assigned the code 'absent'. Students who contributed one or two questions were assigned the code 'low'. The code 'moderate' was assigned to students who contributed between three and four questions. The code 'high' was assigned to students who contributed five questions or more. In this category, the level of participation was assigned on a quantitative measure without an evaluation of the content or the structure of the questions.

SPACE GAME

As I pointed out earlier in this chapter, the two discussions in *Space Game* were coded separately. For each discussion, students who contributed only one idea were assigned the code 'low'. Students who contributed two ideas were assigned the code 'moderate'. Those who contributed three ideas or more were assigned the code 'high'.

Tables 14 and 15 present the mean idea levels for the different group types for Years 4 and 8 respectively.

The mean idea levels generated within the groups were not consistent across the two ages. At Year 4, in general, more ideas were shared in the mixed-gender groups than in the same-gender groups. At Year 8, on average, the 4g groups generated more ideas than the other group types. The 4g groups moved from being the group type that had the lowest mean level of ideas at Year 4 to being the group that had the highest mean level of ideas at Year 8. At both ages, the 3b1g and the 1b3g groups had a higher mean level of ideas than the 2b2g groups.

TABLE 14 YEAR 4: MEAN IDEA LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	1.25	1.28	1.13	1.56	1.29
Language	Brainstorming†	1.90	2.03	2.03	2.00	1.38
Tech	Discussion one	1.41	1.90	1.90	2.10	1.50
	Discussion two	1.00	1.20	1.00	1.20	0.95
AVERAGE		**1.39**	**1.60**	**1.52**	**1.72**	**1.28**

† Same results as for participation.

TABLE 15 YEAR 8: MEAN IDEA LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	1.61	1.33	1.23	1.28	1.63
Language	Brainstorming†	2.10	2.03	1.93	2.15	1.96
Tech	Discussion one	1.36	1.78	1.53	1.80	2.00
	Discussion two	0.71	0.90	0.68	0.98	1.33
AVERAGE		**1.45**	**1.51**	**1.34**	**1.55**	**1.73**

† Same results as for participation.

4 MEAN IDEA LEVELS BY BOYS AND GIRLS IN THE DIFFERENT GROUP TYPES

I also examined the contribution of ideas by boys and girls in the different group types so as to determine whether boys and girls, as subgroups, generate more or fewer ideas in particular group types. The results for the Years 4 and 8 group types are presented in **Tables 16 and 17.**

The overall results relating to the generation of ideas in the various groups indicated different participation patterns for boys and girls at the two ages. At Year 4, girls in the 3b1g groups contributed fewer ideas than did girls in the other group types. Conversely, boys in the 1b3g groups contributed more ideas than did boys in the other group types. Therefore, the issue of being the minority student in a group affected boys and girls differently at Year 4. However, in the Year 8 groups, both the boys' and the girls' groups contributed fewer ideas in the 2b2g groups than in the other group types.

In Discussion One of *Space Game*, boys generated more ideas than girls in all three mixed-gender group types at Year 4 and in two group types at Year 8. The subject area and the content of the activity (improving a board game on space) may have contributed to this disparity. However, it also needs to be pointed out that this disparity occurred in only one of the two discussions that made up the technology task and was not evident in either of the other two tasks.

5 INVOLVEMENT IN ORGANISATION WITHIN THE DIFFERENT GROUP TYPES

The analysis of individual participation concluded with an examination of the involvement of boys and girls in administratively running their group. For this analysis I coded the group members' involvement in organising their group at any stage of the task. The behaviour coded was similar for the different tasks and included raising issues about the procedure to be followed,

TABLE 16 YEAR 4: MEAN IDEA LEVELS BY GENDER AND GROUP TYPE

		4B	3B1G		2B2G		1B3G		4G
		B	B	G	B	G	B	G	G
Science	Discussion	1.25	1.30	1.20	1.05	1.20	1.70	1.50	1.29
Language	Brainstorming†	1.90	2.00	2.10	2.25	1.80	1.80	2.07	1.38
Technology	Discussion one	1.41	1.63	1.00	1.80	1.45	2.00	1.57	1.50
	Discussion two	1.00	0.90	0.60	0.60	0.85	0.80	0.73	0.95
AVERAGE		**1.39**	**1.46**	**1.23**	**1.43**	**1.33**	**1.58**	**1.47**	**1.28**

†Same results as for participation.

TABLE 17 YEAR 8: MEAN IDEA LEVELS BY GENDER AND GROUP TYPE

		4B	3B1G		2B2G		1B3G		4G
		B	B	G	B	G	B	G	G
Science	Discussion	1.61	1.37	1.40	1.15	1.30	1.20	1.30	1.63
Language	Brainstorming†	2.10	1.90	2.40	1.90	1.95	1.80	2.27	1.96
Technology	Discussion one	1.36	1.87	1.50	1.75	1.30	1.80	1.80	2.00
	Discussion two	0.71	0.90	0.90	0.70	0.65	1.10	0.93	1.33
AVERAGE		**1.45**	**1.51**	**1.55**	**1.38**	**1.30**	**1.48**	**1.58**	**1.73**

† Same results as for participation.

distributing the work among the members, keeping track of the ideas discussed/chosen, and making the group aware of issues relating to time and noise.

Students who were not involved in the running of their group were assigned the code 'absent', while those who said something in a one-off situation and appeared to be only marginally concerned were assigned the code 'low'. Students who periodically appeared to be involved in administrative activity were assigned the code 'moderate'. Those who appeared to be in charge most of the time were assigned the code 'high'.

Tables 18 and 19 present the percentages of students involved in organisation at any level within the different group types at Years 4 and 8. Overall, the analysis of organisation at the group level revealed that the percentages of students involved in the different group types were similar at both ages. However, it is worth noting the extent to which this involvement varied across the different tasks within all of the group types.

TABLE 18 YEAR 4: DEGREE OF INVOLVEMENT IN ORGANISATION BY GROUP TYPE

	4B	3B1G	2B2G	1B3G	4G
	%	%	%	%	%
Science	20.8	52.5	50.0	50.0	64.3
Language	40.0	42.5	47.5	37.5	50.0
Technology	75.0	87.5	87.5	82.5	65.0
AVERAGE	**45.3**	**60.8**	**61.7**	**56.7**	**59.8**

TABLE 19 YEAR 8: DEGREE OF INVOLVEMENT IN ORGANISATION BY GROUP TYPE

	4B	3B1G	2B2G	1B3G	4G
	%	%	%	%	%
Science	60.7	55.0	62.5	60.0	45.8
Language	22.5	37.5	42.5	42.5	60.7
Technology	85.7	87.5	80.0	75.0	93.8
AVERAGE	**56.3**	**60.0**	**61.7**	**59.2**	**66.8**

6 INVOLVEMENT OF BOYS AND GIRLS IN ORGANISATION WITHIN THE DIFFERENT GROUP TYPES

Tables 20 and 21 (overleaf) present the percentages of boys and girls (as subgroups) that were involved in organisation within the different group types at Years 4 and 8. The Year 4 results show that the percentages of boys and girls involved in organisation increased as the number of the group members of their gender decreased. Therefore, both boys and girls were most involved in organisation within the group type where they were outnumbered. This was also the case for boys at Year 8 but not for girls.

In all three tasks, for at least one age level, there was an inverse relationship between the percentages of students (boys and/or girls) involved in organisation and their mean participation levels in the different group types. Therefore, when students became highly involved in organisation, their on-task participation tended to be low and vice versa. This observation was only possible because on-task participation and involvement in organisation were coded separately.

In two of the three tasks (*Separating Mixtures* and *Question Time*) there were more girls than boys involved in organisation at both age levels. This pattern was overturned in *Space Game* where, overall, more boys than girls were involved in organisation at both age groups. These results suggest that involvement in organisation generally related more to the subject area and the nature of the task than to the age level of the students or their gender characteristics.

SUMMARY

The analysis of the three tasks, *Separating Mixtures, Question Time* and *Space Game,* showed that participation in the different group types did not remain consistent across the different activities that made up any one task. Therefore, to present an accurate picture, the analysis of each activity needed to be presented separately. Overall, the video analysis did not identify any group types that had higher participation levels consistently across the three tasks. However, especially at the Year 4 level, 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 his or her gender group working in other group settings. At the same time, the minority students in the 3b1g and the 1b3g groups tended to become highly involved in the organisation of their group. As educators we need to keep a look out for these two phenomena in order to ensure that the different group members benefit equally from group activities.

TABLE 20 YEAR 4: DEGREE OF INVOLVEMENT IN ORGANISATION BY GENDER AND GROUP TYPE

	4B	3B1G		2B2G		1B3G		4G
	B%	B%	G%	B%	G%	B%	G%	G%
Science	20.8	43.3	80.0	45.0	55.0	50.0	50.0	64.3
Language	40.0	33.3	70.0	40.0	55.0	50.0	33.3	50.0
Technology	75.0	93.3	70.0	90.0	85.0	80.0	83.3	65.0
AVERAGE	**45.3**	**56.6**	**73.3**	**58.3**	**65.0**	**60.0**	**55.5**	**59.8**

TABLE 21 YEAR 8: DEGREE OF INVOLVEMENT IN ORGANISATION BY GENDER AND GROUP TYPE

	4B	3B1G		2B2G		1B3G		4G
	B(%)	B(%)	G(%)	B(%)	G(%)	B(%)	G(%)	G(%)
Science	60.7	3.3	60.0	50.0	75.0	60.0	60.0	45.8
Language	22.5	33.3	50.0	25.0	60.0	40.0	43.3	60.7
Technology	85.7	90.0	80.0	85.0	75.0	90.0	70.0	93.8
AVERAGE	**56.3**	**58.9**	**63.3**	**53.3**	**70.0**	**63.3**	**57.8**	**66.8**

4

SOCIAL PROCESSES IN THE FIVE GROUP TYPES

This chapter focuses on three social processes — interaction, co-operation and conflict. It aims to answer the second research question, namely, the relationship between the group gender composition and these processes within groups. It is important, while reading this chapter, to keep in mind that in the analysis of the group processes, the focus shifts from the individual group members to the group as one unit.

The results relating to the three processes are reported separately here. Section One reports on interaction, Section Two on co-operation and Section Three on conflict. In each section, I first define the specific process and explain its role in group work by drawing on our current understanding derived from research. I then describe the coding process involved in the video analysis and present the results for the five group types at Years 4 and 8 separately. In this way it is possible to evaluate whether particular processes were more or less common in certain group types and whether this was the case in one of or in both the two age groups.

SECTION ONE: INTERACTION

DEFINITION AND ROLE

Effective group work involves more than contributions from individual students. For the collaborative process to be effective, these contributions must interact with those of other students. Salomon and Globerson (1989, 93) explain that

> the very fact that a team not an individual learner is involved implies that the interaction among group members is not just unrelated questions and answers, queries and responses and individuals' cognitive processes. A team is a social system ... behaviour and cognition become interdependent.

Webb (1982c, 1989, 1991) and Webb and Kenderski (1985) have analysed the amount of interaction that occurs during group work as well as different aspects of interactive verbal behaviour, such as giving and receiving explanations and asking for and giving different kinds of help. According to Webb (1991, 366), our current understanding of the nature of students' verbal interactions during group work is still limited, for 'only a minority of research has examined the kinds of task-related verbal interaction that occurs when students work together'. She argues that insight into such interactive behaviour is required if we are to achieve a fuller understanding of the effects of group work on students' achievement. In particular, such insight may shed light on the conflicting results of studies that have addressed the relationship between interaction and achievement.

CODING THE INTERACTION LEVELS

In this study, interaction was investigated on two occasions: during the discussion of the science task and during the question choice activity of the language task. In both cases, I defined verbal interaction as the extent to which the group members talked *to one another*. This category included behaviours like sharing ideas and information, giving suggestions and opinions about ideas, questioning ideas, and other involvement in the discussion. I coded the level of the group's verbal interaction using the ratings 'absent', 'low', 'moderate' or 'high'. These ratings were established for each activity separately, using video exemplars to differentiate among the levels.

RESULTS

The interaction results for the Years 4 and 8 group types are presented in **Tables 22 and 23** respectively. These tables show that the interaction results were different for the two age groups. At Year 4, the 4b groups had a higher mean interaction level than the other group types. At Year 8, two matters are noteworthy: in general, the 2b2g groups had a lower interaction mean than the other group types, and the same-gender groups had a higher interaction mean than the mixed-gender groups. However, the results of the two activities taken separately indicated that higher interaction in the same-gender groups occurred in only one of the activities at both Years 4 and 8. Overall, across the two activities in the two age groups, the 4b groups were observed to interact the most.

TABLE 22 YEAR 4: MEAN INTERACTION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	2.17	1.80	2.10	1.90	1.71
Language	Choice	2.00	1.90	1.80	1.70	2.00
AVERAGE		**2.09**	**1.85**	**1.95**	**1.80**	**1.86**

TABLE 23 YEAR 8: MEAN INTERACTION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	2.29	1.80	1.60	1.70	2.17
Language	Choice	2.10	1.90	1.40	2.00	1.90
AVERAGE		**2.20**	**1.85**	**1.50**	**1.85**	**2.04**

SECTION TWO: CO-OPERATION

DEFINITION AND ROLE

The nature and extent of students' co-operative behaviour is another issue that has been the focus of discussions on the social processes of group work. Various researchers (e.g., Cohen, 1986; Hall, 1994; Hertz-Lazarowitz, 1992; Johnson & Johnson, 1992; Kagan, 1992, and Solomon, Watson, Schaps, Battistich & Solomon, 1990, both cited in Webb, 1994) have defined co-operative behaviour in a number of ways, such as students'

- responsiveness to the needs of the group and to the problems of their group members
- awareness of the nature of collective decision-making
- willingness to help one another
- understanding and appreciation of others
- ability to provide effective feedback, support and encouragement
- awareness of the importance of turn-taking.

Hertz-Lazarowitz (1992) and Johnson and Johnson (1992) suggest that the development of these interpersonal skills is a necessary condition of group work.

CODING THE CO-OPERATION LEVELS

I coded the co-operation level of each group using the ratings 'absent', 'low', 'moderate' or 'high'. These ratings were established for each activity separately using video exemplars to differentiate among the levels. Co-operation was not limited to verbal contributions towards the end of the task but also included the group members' attitudes towards one another and towards what they said, and their concern or otherwise for the involvement of the other team members. Although co-operation was not limited to verbal contributions, I perceived any student who did not participate at all as having not co-operated.

SEPARATING MIXTURES

In the science task *Separating Mixtures,* the groups' co-operation levels were coded separately for the discussion and the experiment, and therefore the group members' efforts and ability to work together during a verbal activity (the discussion) and a physical one (the experiment) were evaluated separately.

QUESTION TIME

In *Question Time*, co-operation was coded during the question choice activity and during the reporting and justification activity. During the former, co-operation meant the students' efforts and ability to work together as they went through the process of choice. During the latter, the decision on the groups' co-operation level was taken globally, as the different group members' participation was not required at all times. However, the students' willingness to add to what others had said and their efforts to help other group members (for instance, when a student appeared unsure of which question to report, why a particular question had been chosen or how to read a word) were indications of co-operation. On the other hand, refusal to participate and unwillingness to help out other group members at any particular point were indications of non-co-operation.

SPACE GAME

In the technology task *Space Game,* each group's co-operation levels in the two discussions were coded separately. In each case, the coding of this category was limited to the discussion part of the activity when the children were left to work on their own.

RESULTS

The mean co-operation levels of the five group types are presented separately for the two age levels in **Tables 24 and 25**. Generally, the amount of co-operation varied in the different activities that made up the tasks. Overall, however, the older groups co-operated more than the younger ones in all of the group types. It is worth pointing out that the 4g groups had the highest co-operation mean at both Years 4 and 8, but there was no indication that the same-gender groups co-operated more or less than the mixed-gender groups. The 3b1g groups stood out as the groups that experienced most difficulty working together in several activities at Year 4.

TABLE 24 YEAR 4: MEAN CO-OPERATION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	2.50	1.50	2.10	2.20	1.86
	Experiment	2.17	1.80	2.30	2.40	2.57
Language	Choice	2.20	2.10	1.80	2.30	2.00
	Justification	1.80	1.80	2.00	1.80	2.50
Tech	Discussion one	2.00	1.70	2.10	2.00	2.00
	Discussion two	1.50	1.70	1.80	1.70	2.00
AVERAGE		**2.03**	**1.77**	**2.02**	**2.07**	**2.16**

TABLE 25 YEAR 8: MEAN CO-OPERATION LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	2.00	2.00	2.30	2.10	2.33
	Experiment	2.57	2.70	2.20	2.70	2.83
Language	Choice	2.30	2.00	2.00	2.20	2.30
	Justification	2.30	2.00	2.10	2.00	2.00
Tech	Discussion one	2.14	2.40	2.60	2.60	3.00
	Discussion two	1.83	1.90	1.70	1.90	3.00
AVERAGE		**2.19**	**2.17**	**2.15**	**2.25**	**2.58**

SECTION THREE: CONFLICT

DEFINITION AND ROLE

Conflict is another social process in group work that has provoked much discussion. Although it is widely acknowledged that conflict (defined as incompatibility among students' behaviours or goals (Shantz, 1987)) is an inevitable aspect of collaborative group work, there is less consensus on the extent to which conflict makes group activity more effective.

Johnson and Johnson (1994) have been among the most vocal advocates of the benefits of conflict in group work. Drawing on Piaget's notions of cognitive conflict, they argue that conflict is a necessary component of successful group work and that co-operative behaviours, such as those listed in the previous section, do not, in themselves, ensure that a group will be maximally productive. However, an earlier study by Lindow, Wilkinson and Peters (1985) did not find empirical support for this assertion. In their study, students who participated more frequently in verbal disagreements during group work performed only marginally better on a subsequent achievement test than students who had less verbal conflict.

Bearison, Magzamen and Filardo (1986), cited in Webb & Palinscar (1996), identified a complex relationship between conflict and learning. They found that dyads that engaged in infrequent or very frequent verbal disagreements gained less *on an achievement test* than those that engaged in a moderate amount of verbal disagreement. Webb and Palinscar (1996) suggest that infrequent conflict may reflect suppression of disagreements whereas too much conflict may prevent children from seeking new information to resolve their disagreements.

Sharan (1990, 297) remains cautious about the extent to which conflict during group work necessarily promotes higher levels of critical thinking in students, concluding that 'the features and conditions of cognitive controversy that can stimulate critical thinking remain to be studied'.

CODING THE CONFLICT LEVELS

My interpretation of conflict implied interpersonal controversy and overt opposition by one person to another person's actions or statements. This category included behaviours such as confrontation and arguments about turn-taking and procedures to be followed. Negative reactions to other group members' ideas or choices were not considered as conflicts unless they developed into arguments.

Conflict was coded using the ratings 'not at all', 'rarely', 'moderately', 'lots'. This coding scheme included both the number and the intensity of the conflicts, and video exemplars were used as indicators of the different levels for the various activities.

RESULTS

The results of the analysis of conflict in the five group types are presented in **Tables 26 and 27** for Years 4 and 8 groups respectively. The overall results relating to conflict suggested that this characteristic was present almost evenly in all of the group types at both age levels. **Tables 26 and 27** show that conflict was more common in the Year 4 groups than in the Year 8 groups. This finding was reflected in the conflict means of the different group types as well as in the analysis of the percentages of groups displaying conflict in the different activities (see **Tables 28 and 29,** overleaf). These findings suggest that there is a relationship between the amount of conflict occurring in groups generally and the age of the children.

The analysis of conflict also showed that the percentages of groups displaying conflict differed in terms of the various activities that made up each of the three tasks. In *Separating Mixtures* there was more conflict during the discussion than during the experiment, and in *Question Time* there was more conflict during the question choice activity than during the reporting/justification activity. In *Space Game* the most conflict occurred in the game, and there was more conflict in Discussion One than in Discussion Two. These findings suggest that the amount of conflict varies within activities that make up one task.

TABLE 26 YEAR 4: MEAN CONFLICT LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	0.83	0.90	0.70	1.00	1.14
	Experiment	1.17	1.20	0.90	0.90	0.71
Language	Choice	0.60	0.80	1.00	0.50	2.00
	Justification	0.40	0.10	0.40	0.10	0.50
Tech	Game	0.86	1.30	0.70	0.40	0.40
	Discussion one	0.63	0.50	0.80	0.50	0.60
	Discussion two	0.17	0.40	0.20	0.10	0.60
AVERAGE		**0.67**	**0.74**	**0.67**	**0.50**	**0.85**

TABLE 27 YEAR 8: MEAN CONFLICT LEVELS BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	0.71	1.00	0.60	0.90	1.00
	Experiment	0.29	0.10	0.30	0.20	0.00
Language	Choice	0.20	0.10	0.10	0.50	0.00
	Justification	0.00	0.00	0.10	0.40	0.00
Tech	Game	0.71	0.40	0.40	0.20	0.75
	Discussion one	0.43	0.10	0.40	0.30	0.50
	Discussion two	0.00	0.00	0.10	0.50	0.00
AVERAGE		**0.33**	**0.24**	**0.29**	**0.43**	**0.32**

The analysis of conflict also showed that the amount of conflict in particular group types differed according to the nature of the activity. 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 (e.g., carrying out an experiment or playing a board game). These observations show that age and the nature of the activity have more influence on conflict than has group gender composition.

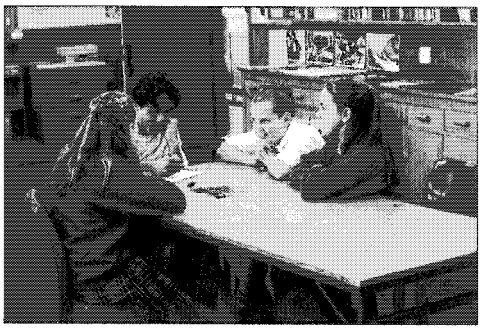
Finally, the analysis of conflict highlighted three other findings. First, there were no differences between boys and girls in terms of how much they initiated and participated in conflicts. Second, the majority of the conflicts that took place in mixed-gender groups were mixed-gender conflicts. And, third, in many cases conflicts included more than two group members.

TABLE 28 YEAR 4: PERCENTAGES OF GROUPS DISPLAYING CONFLICT

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	67	60	60	80	86
	Experiment	83	70	60	50	57
Language	Choice	40	50	70	50	100
	Justification	40	10	30	10	50
Tech	Game	86	70	60	40	40
	Discussion one	63	40	60	40	40
	Discussion two	17	30	20	10	40
AVERAGE		**57**	**47**	**51**	**40**	**59**

TABLE 29 YEAR 8: PERCENTAGES OF GROUPS DISPLAYING CONFLICT

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	43	70	60	70	83
	Experiment	29	10	20	10	0
Language	Choice	20	10	10	20	0
	Justification	0	0	10	10	0
Tech	Game	71	40	40	20	50
	Discussion one	43	10	40	30	50
	Discussion two	0	0	10	30	0
AVERAGE		**29**	**20**	**27**	**27**	**26**



SUMMARY

Overall, the analyses of the relationship between group gender composition and levels of interaction, co-operation and conflict showed that these experiences were relatively similar across the different group types. Certainly, there were no clear divisions between the experiences in the same- and the mixed-gender groups. In general, it was one group type that stood out in the different analyses: for example, the 4b groups were observed to interact the most while the 4g groups were observed to co-operate the most.

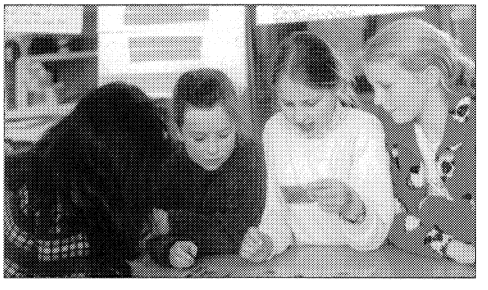
The analysis of conflict did not find differentiation among the five group types generally. However, age and the nature of the activity affected all of the group types, sometimes in a similar manner (e.g., the Year 4 group types experienced more conflict than the Year 8 group types) and sometimes in a different manner (conflict was higher in the 4g groups during verbal activities, and it was higher in the 4b groups during psychomotor activities).

5

JOINT PRODUCTS OF THE FIVE GROUP TYPES

In this chapter, the scores of the joint products (i.e., the collective achievement of each group) on the three tasks were analyzed in order to answer the third research question, which dealt with the relationship between the group gender composition and achievement on the group tasks. I first explain the process used to obtain the scores for the joint products. I then present the results of the analysis of the joint products for the different group types and follow this with an examination of the relationships between the group processes reported in the previous chapter and the joint products.

It should be explained at this point that, in the NEMP, all group tasks are video-recorded, and the data are couriered to the Educational Assessment Research Unit in Dunedin. In January of each year, teacher-markers from throughout New Zealand meet in Dunedin to mark these tasks. They first work together to set the benchmarks for assessing the children's achievement on each task and then work in pairs to mark the work of a sample of groups. The same marking criteria are used for tasks that are carried out at both age levels, with Years 4 and 8 groups intermingled in the marking process. In this chapter, the analysis is based on marks given by the teacher-markers.



PROBE STUDY REPORT

THE CODING PROCESS

SEPARATING MIXTURES

The two activities that made up *Separating Mixtures* were marked independently. Each group was given a discussion score and an experiment score, made up from marks for specific aspects of the group task. The marking schedule used for scoring this task is presented in the Appendix. (Note that the activity that I have called 'Discussion' in my study is listed as 'Planning' on the marking schedule.)

QUESTION TIME

The groups were given one overall score for *Question Time*. The marking schedule used for scoring this task is presented in the Appendix.

SPACE GAME

For *Space Game*, each group obtained a separate score for each of the two discussions. No score was given for the game. The marking schedule used for scoring this task is presented in the Appendix.

RESULTS

THE JOINT PRODUCTS

Given that the maximum scores for the various activities were different, I decided to present in this analysis the mean group *z*-scores instead of the mean group raw scores, so as to make comparisons across the three tasks possible. The results for the Years 4 and 8 group types are presented in **Tables 30 and 31**.

The overall mean group *z*-scores for the two ages indicated that only one group type (the 1b3g groups) had a mean score that was above the overall mean score at both Years 4 and 8. In other words, the 1b3g groups were the only ones to score higher than the average mean score at both ages. The 4b groups, on the contrary, had a *z*-score that was below the overall mean score at both ages. The scores for the other group types were inconsistent, with the 2b2g groups having a positive *z*-score at Year 4 and a negative *z*-score at Year 8, and the 4g and the 3b1g groups exhibiting the opposite pattern.

It is worth noting that at Year 4 both of the same-gender group types had an overall negative *z*-score. At this age level, the achievement in the same-gender groups was lower than in the mixed-gender groups. In the Year 8 groups, there was no pattern that clearly separated the same- and the mixed-gender groups.

The mean *z*-scores in the different group types appeared to be influenced at times by the following: subject area (e.g., the 4g groups had a positive mean *z*-score on the language task at both age groups); the nature of the task (e.g., at Year 4, the 4g groups had a positive mean *z*-score on the language task and a negative mean *z*-score for both technology task discussions); and the age of the group members (e.g., the 4g groups had the highest mean *z*-score for the science task experiment at Year 4 and a negative mean *z*-score for the same activity at Year 8).

TABLE 30 YEAR 4: MEAN GROUP Z-SCORES BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	+0.66	0.00	+1.27	−1.16	−0.78
	Experiment	−0.47	−1.23	−0.37	+0.98	+1.08
Language	Overall	−1.40	−0.39	0.00	+1.28	+0.51
Tech	Discussion one	+0.50	+0.66	+0.66	−0.13	−1.69
	Discussion two	−0.02	+0.77	+0.16	+0.77	−1.67
AVERAGE		**−0.15**	**−0.04**	**+0.34**	**+0.35**	**−0.51**

TABLE 31 YEAR 8: MEAN GROUP Z SCORES BY GROUP TYPE

		4B	3B1G	2B2G	1B3G	4G
Science	Discussion	+0.43	+1.15	−0.65	+0.43	−1.37
	Experiment	−0.08	+1.48	−0.43	+0.27	−1.24
Language	Overall	−1.23	+0.70	−0.53	−0.23	+1.29
Tech	Discussion one	−1.59	−0.35	+0.84	+0.68	+0.42
	Discussion two	−0.55	−0.62	−0.62	+0.06	+1.71
AVERAGE		**−0.60**	**+0.47**	**−0.28**	**+0.24**	**+0.16**

RELATIONSHIPS BETWEEN THE GROUP PROCESSES AND THE JOINT PRODUCTS

In this section, I examine the relationships between the group processes and the joint products for the three tasks at the two age levels. In these analyses, the data were correlated using the results of all the individual groups. However, in the following tables I present the group means as an indication of the results for the different group types. It is important to point out that, in these analyses, a correlation of 0.4 or higher is statistically significant ($p < 0.01$).

SEPARATING MIXTURES

The results for *Separating Mixtures* at Year 4 indicated the following relationships within groups (see **Table 32**). High relationships were evident between:

a) the level of interaction and the joint product in the discussion ($r = 0.74$)

b) the level of co-operation and the joint product in the experiment ($r = 0.92$)

c) the level of conflict and the joint product in the discussion ($r = -0.90$) and the experiment ($r = -0.85$).

TABLE32 SCIENCE TASK, YEAR 4: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

		4B	3B1G	2B2G	1B3G	4G
N		6	10	10	10	7
Discussion	Mean product score	5.67	5.10	6.20	4.10	4.43
	Mean interaction level	2.17	1.80	2.10	1.90	1.71
	Mean co-op. level	2.50	1.50	2.10	2.20	1.86
	Mean conflict	1.17	1.20	0.90	0.90	0.71
Experiment	Mean product level	2.67	2.44	2.70	3.11	3.14
	Mean co-op. level	2.17	1.80	2.30	2.40	2.57
	Mean conflict level	0.83	0.9	0.70	1.00	1.14

TABLE33 SCIENCE TASK, YEAR 8: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

		4B	3B1G	2B2G	1B3G	4G
N		7	10	10	10	6
Discussion	Mean product score	8.00	8.20	7.70	8	7.50
	Mean interaction level	2.29	1.80	1.60	1.70	2.17
	Mean co-op. level	2.00	2.00	2.30	2.10	2.33
	Mean conflict level	0.71	1.00	0.60	0.90	1.00
Experiment	Mean product level	4.50	5.40	4.30	4.70	3.83
	Mean co-op. level	2.57	2.70	2.20	2.70	2.83
	Mean conflict level	0.29	0.10	0.30	0.20	0.00

A low relationship was evident between:

d) the level of co-operation and the joint product in the discussion ($r = 0.22$).

In essence, the main points to emerge from this analysis were the following:

1 The more the group members talked to one another during the discussion, the higher the group scored on the joint product.

2 The more the group members were able to work together during the experiment, the higher the group scored on the joint product.

3 The more conflict a group experienced, the lower its joint product scores for both the discussion and the experiment.

The results for *Separating Mixtures* at Year 8 indicated the following relationships within groups (see **Table 33**). A high relationship was evident between:

a) the level of co-operation and the joint product in the discussion ($r = -0.94$).

Low relationships were evident between:

b) the level of co-operation and the joint product in the experiment ($r = 0.06$)

c) the level of interaction and the joint product in the discussion ($r = -0.17$)

d) the level of conflict and the joint product in the discussion ($r = -0.13$) and the experiment ($r = -0.10$).

Essentially, the analysis showed that the more the group members co-operated during the discussion, the lower they scored on the joint product.

QUESTION TIME

The results for *Question Time* at Year 4 indicated the following relationships within groups (see **Table 34**). A high relationship was evident between:

a) the level of interaction during the question choice and the joint product (r = -0.80).

Moderate relationships were evident between:

b) the level of ideas generated within a group and the group's joint product (r = 0.49)

c) a group's ability to justify its choices and its joint product (r = -0.38)

d) a group's overall co-operation level and its joint product (r = -0.40).

A low relationship was apparent between:

e) a group's overall conflict level and its joint product (r = -0.01).

TABLE34 LANGUAGE TASK, YEAR 4: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

	4B	3B1G	2B2G	1B3G	4G
N	4	7	10	10	2
Mean product score	7.50	8.29	8.60	9.60	9.00
Mean idea level	2.00	2.29	2.40	2.40	2.00
Mean interaction level	2.25	1.86	1.80	1.70	2.00
Mean justification level	2.25	2.14	2.00	1.80	2.50
Mean co-op. level	4.5	4.15	3.80	4.10	4.50
Mean conflict level	1.25	0.85	1.44	0.60	2.50

TABLE35 LANGUAGE TASK, YEAR 8: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

	4B	3B1G	2B2G	1B3G	4G
N	9	10	9	9	7
Mean product score	9.89	11.4	10.44	10.67	11.86
Mean idea level	2.67	2.40	2.33	2.44	2.29
Mean interaction level	2.11	1.90	1.44	1.89	1.86
Mean justification level	2.56	2.10	2.00	2.14	2.43
Mean co-op. level	4.66	4.00	4.00	4.25	4.29
Mean conflict level	0.11	0.1	0.31	1.19	0

The Question Time data at Year 4 yielded one 'expected' relationship, namely, the more ideas the group members shared, the higher the group scored on the joint product. The other results were not so predictable.

The results for *Question Time* at Year 8 indicated the following relationships within groups (see **Table 35**). A high relationship was evident between:

a) the level of ideas generated and the joint product (r = -0.75).

A moderate relationship was apparent between:

b) the overall mean co-operation level and the joint product (r = -0.46).

Low relationships existed between:

c) the level of interaction during the question choice and the joint product (r = -0.07)

d) the groups' ability to justify their choices and their joint products (r = -0.10)

e) the level of conflict and the joint product (r = -0.26).

The results for the Year 8 groups were not consistent with those for the Year 4 groups. At Year 8, groups generating more ideas tended to get lower joint product scores, whereas at Year 4 groups generating more ideas tended to get higher joint product scores. Moreover, at Year 8 there was a low relationship between a group's interaction level and its joint product, whereas at Year 4 there was a high relationship between these two variables.

SPACE GAME

The results for *Space Game* at Year 4 indicated the following relationships within groups (see **Table 36**). High relationships were evident between:

a) the level of ideas generated in a group and the joint product in both discussions (for Discussion One $r = -0.93$ and for Discussion Two $r = -0.72$)

b) the level of conflict and the joint product in Discussion Two ($r = -0.72$).

A moderate relationship existed between:

c) the level of co-operation and the joint product in Discussion Two ($r = -0.66$).

There were low relationships between:

d) the level of conflict and the joint product in Discussion One ($r = 0.19$)

e) the level of co-operation and the joint product in Discussion One ($r = -0.22$).

TABLE36 TECHNOLOGY TASK, YEAR 4: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

		4B	3B1G	2B2G	1B3G	4G
N		8	10	10	10	5
Discussion one	Mean product score	1.88	1.90	1.90	1.80	1.60
	Mean idea level	2.13	1.90	1.90	2.10	2.20
	Mean co-op. level	2.00	1.70	2.10	2.00	2.00
	Mean conflict level	0.63	0.50	0.80	0.50	0.60
Discussion two	Mean product score	1.67	1.80	1.70	1.80	1.40
	Mean idea level	1.33	1.20	1.00	1.20	1.60
	Mean co-op. level	1.50	1.70	1.80	1.70	2.00
	Mean conflict level	0.17	0.40	0.20	0.10	0.60

TABLE37 TECHNOLOGY TASK, YEAR 8: RELATIONSHIPS BETWEEN JOINT PRODUCTS AND PROCESSES

GROUP TYPES		4B	3B1G	2B2G	1B3G	4G
N		7	10	10	10	4
Discussion one	Mean product score	1.86	2.10	2.33	2.30	2.25
	Mean idea level	1.71	2.30	1.90	2.60	2.50
	Mean co-op. level	2.14	2.40	2.60	2.60	3.00
	Mean conflict level	0.43	0.10	0.40	0.30	0.50
Discussion two	Mean product score	2.33	2.30	2.30	2.60	3.33
	Mean idea level	1.00	1.40	1.20	1.30	1.00
	Mean co-op. level	1.83	1.90	1.70	1.90	3.00
	Mean conflict level	0.00	0.00	0.10	0.50	0.00

The main points to emerge, then, from this analysis were the following:

1 The more that ideas were shared in a group, the lower the joint product score. (This finding suggests that in such groups the group members did not develop a discussion as such but instead volunteered ideas in a parallel manner.)

2 The more conflict the group experienced, the lower the group's score on the joint product for one of the discussions.

3 The groups' co-operation levels were not positively correlated to their joint products.

The results for *Space Game* at Year 8 indicated the following relationships within groups (see **Table 37**). High relationships were apparent between:

a) the level of co-operation and the joint product in both discussions (for Discussion One $r = 0.77$ and for Discussion Two $r = 0.97$)

b) the level of ideas generated in the group and the joint product in Discussion Two ($r = -0.91$).

A moderate relationship existed between:

c) the level of ideas within the group and the joint product in Discussion One ($r = 0.58$).

Low relationships were evident between:

d) the level of conflict within the group and the joint product in both Discussion One (r = 0.05) and Discussion Two (r = -0.03).

The main points that emerged from this analysis, therefore, were the following:

1 The more the groups were able to work together, the higher their joint product score.

2 The more the group members were able to talk to one another during the discussion, the higher their joint product score.

3 Conversely, the more ideas generated in the group, the lower the group's achievement on the joint product. (It is possible that points 2 and 3 together indicate that groups which generated many ideas were the ones where group members did not talk much to each other and hence did not develop a discussion.)

4 There was little relationship between the amount of conflict in the groups and the joint product scores.

SUMMARY

Analyses of the relationship between the group processes and the joint products in the different group types 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 the joint product.

The inconsistent relationships between the co-operation levels and the joint products led me to conclude that although the development of interpersonal/co-operative skills has been acknowledged as a necessary condition for group work (see, for example, Cohen, 1986; Hall, 1994), its presence does not necessarily relate to achievement on the group task, perhaps in the same way that effort does not always relate to achievement.

The results for the relationship between interaction and/or generation of ideas and the joint product suggest that teachers need to make explicit to students whether one or the other is the priority in a particular activity.

6

CHILDREN'S PERSPECTIVES ON GROUP GENDER COMPOSITIONS

In recent years, several social scientists (e.g., Brannen & Brien, 1995; Qvortrup, 1990; Smith, 1995) have stressed the need to incorporate children's views and perspectives into research that involves them.

This new model of research supports 'the social child model' (Prout & James, 1990), which sees children as competent and skilful in providing their views and perspectives on their reality. This model is a reaction against the dominant discourse about the relationships between children and society (Moss & Petrie, 1997), where 'children [are]... seen as the objects of the academic gaze and not recognised as social actors in their own right' (Smith, 1998, 73). According to Prout and James (1990, 8), within the context of this new model of research, 'children are and must be seen as active in the construction and determination of their social lives ....they are not just the passive subjects of social structures and processes.'

Smith (1998, 69) suggests that 'by listening to children's stories, we recognise them as people in their own right at the present moment in time'. Qvortrup (1990, 11) explains that children's perspectives are needed in addition to and not instead of other research approaches:

> 'children's perspectives can be incorporated into research but this does not mean abandonment of observational approaches ... relevant quantitative information about children's lives is also urgently needed.

Similarly, Smith (1998, 17) stresses the need to incorporate children's perspectives as one dimension in research involving them:

> I am not arguing that ... their perspective is the only one, merely that we need to provide opportunities for children to express their views, listen to them respectfully, take them into account.'

During my exploration of the experiences of boys and girls in groups with different gender compositions, I obtained direct information from the children about their views of these experiences by means of a questionnaire and an interview. In this chapter, I first report on the students' views on working in groups with different gender compositions and then report on their evaluations of the NEMP group experiences. My goal here is to answer my last research question, which related to whether the children's evaluation of a particular group task varied in the different group types and whether it was affected by the nature of the task.

METHOD

All students who participated in the NEMP in 1996 completed the questionnaire, and their responses to general questions regarding involvement in the different types of groups are reported here. The questionnaire responses of all the four-member groups that completed *Space Game* and *Question Time* are then used to present the students' evaluation of the specific group work activities.

The students interviewed were the 23 eight- to nine-year-olds who were randomly selected from two schools to participate in the NEMP Year 4 assessment in Dunedin in 1996. Two techniques were used in the one-to-one interviews: a structured interview and a stimulated recall interview. However, only the former is reported on in this chapter. The interview codes for the children from School 1 are A1-4, B1-4 and C1-4. For School 2, they are AA1-4, BB1-4 and CC1-4. The numbers following these codes in the quotations refer to the line numbers of the children's transcripts.

CONDUCTING THE ANALYSIS

Statistical tests were carried out on the questionnaire data in two ways. I first tested for gender differences within each group type. Using data obtained from all the respondents, I checked for statistically significant differences in the means for boys and the means for girls in the 3b1g, 2b2g and 1b3g groups. I also looked for any significant differences between the 4b and 4g groups. The second test involved within-gender differences across the group types. To check for statistically significant differences between the means for boys, I used data from the 4b, 3b1g, 2b2g and 1b3g groups. For girls, I used data from the 4g, 1b3g, 2b2g and 3b1g groups.

† Student code, transcript line number.

RESULTS

CHILDREN'S ATTITUDES TOWARD SAME-GENDER AND MIXED-GENDER GROUPS

By means of the questionnaire, I asked all of the children participating in the NEMP project how they felt about working in different group types (with boys, with girls, and with both boys and girls). I then tried, using a structured interview to question a sub-sample of the students, to explore some of the reasons behind these attitudes.

In the questionnaire, the children rated how much they enjoyed working in the different group types, using the faces five-point rating scale shown below. Here, 5 and 4 portray happy faces, 3 is neutral and 2 and 1 portray unhappy faces.

5 4 3 2 1

At Year 4, as is evident from **Table 38**, both boys and girls gained the most enjoyment from working in same-gender groups. In the interviews, the boys explained that working with other boys was fun. Boys were nice(r) to one another, understood one another, liked the same kinds of things, co-operated and always found something to do. They seemed aware of a certain fraternity amongst boys, as these examples illustrate:

Chris: They understand what you are talking about. (C2, 72†)

Daniel: Most of the time we co-operate and we know what to talk about and we never get bored. (B4, 104†)

Mick: It's fun and people are nicer to you. (C4, 79†)

The girls explained that they enjoyed being with other girls and that working with them was fun and easy because they communicated better, shared ideas, agreed on what they had to do, co-operated and helped one another. The following two examples illustrate these points.

Anna: It's fun because we all co-operate on one thing. (C1, 80†)

Hayley: It's really easy because girls can communicate better than boys ... they can be easy and help you. (B2, 98, 107†)

The questionnaire results also showed that the Year 4 children gained greater enjoyment from working in balanced mixed-gender groups than in groups with members of the other gender only. In the interviews, the children explained that they felt comfortable in balanced mixed groups because they had the support and help of members of their own gender:

Lisa: No-one's feeling the odd one out. (B1, 161†)

Kylie: You have girls in the team to help you. (AA2, 143†)

Daniel: Well, I feel fine because there's people that I do like, and if the girls start to yell at me and I get a little unhappy I can just talk to my friends and I get happy again.

Grace: You mean to the other boys in the group?

Daniel: Yes. (B4, 122–128†)

All of the children that I interviewed (with the exception of one girl) thought that it was a good idea for boys and girls to work together because in this way they could learn about one another and possibly become friends. The children therefore showed a general awareness of the benefits of balanced mixed-gender groups and they themselves did not object to being in such groups.

Working in a group with members of the other gender only was the least attractive option for both boys and girls at Year 4. In the interview, the children talked about discomfort, insecurity, peer pressure and a lack of mixed-gender friendships:

Meg: I won't feel that good because I'll be the only one who's a girl and I won't be able to feel comfortable. (AA1, 115–16†)

Scott: [Laughs, then says] you feel a bit lonely. (BB4, 141†)

Danielle: Boys are friends, and they would go, 'You are a girl, I can't be friends with you.' (A2, 129†)

Lisa: [I'd rather be with girls] so that I could be with some people which I know. (B1, 149†)

Mick: Well, people tease me and the boys tease me and they don't like me anymore. (C4, 95†)

At Year 8 (see **Table 38**), boys and girls responded equally favourably to same-gender and balanced mixed- gender groups, but still felt less positive about working with members of the other gender only. For both genders, the mean for working in balanced mixed-gender groups was higher at Year 8 than at Year 4. However, at both age levels, the mean for girls was significantly higher than the mean for boys in this group type. Girls enjoyed working in balanced gender-mixed groups more than did boys.

TABLE 38 THE EXTENT TO WHICH BOYS AND GIRLS ENJOYED WORKING IN SAME- AND MIXED-GENDER GROUPS ACCORDING TO 5-POINT FACES RATING SCALE

		IN SAME-GENDER GROUPS X RATING	IN GENDER-BALANCED GROUPS X RATING	WITH MEMBERS OF THE OTHER GENDER ONLY X RATING
Year 4	Boys	4.72	4.09†	3.02
	Girls	4.79	4.37†	3.25
Year 8	Boys	4.63	4.58†	4.19
	Girls	4.72	4.76†	4.27

NOTE: Gender differences within group types † = $p < 0.001$.

CHILDREN'S EVALUATION OF EXPERIENCES IN GROUPS WITH DIFFERENT GENDER COMPOSITIONS

I was also interested in finding out whether the children's evaluation of a particular group experience varied according to the gender composition of their group. I investigated this by means of a post-task individual questionnaire, which was completed

† Student code, transcript line number.

TABLE 39 THE EXTENT TO WHICH CHILDREN LIKED BEING IN THEIR NEMP GROUP ACCORDING TO 5-POINT FACES RATING SCALE

		N	%	😊	%	🙂	%	😐	%	🙁	%	☹	%	NA†	%
QUESTION TIME GROUPS															
Year 4	Boys	208	100	147	70.67	37	17.79	9	4.33	5	2.40	9	4.33	1	0.48
	Girls	192	100	142	73.96	28	14.58	11	5.73	5	2.60	3	1.56	3	1.56
Year 8	Boys	193	100	107	55.44	60	31.09	18	9.33	3	1.55	2	1.04	3	1.55
	Girls	163	100	103	63.19	46	28.22	9	5.52	2	1.23	1	0.61	2	1.23
SPACE GAME GROUPS															
Year 4	Boys	219	100	159	72.60	35	15.98	11	5.02	4	1.83	6	2.74	4	1.83
	Girls	189	100	146	77.25	26	13.76	13	6.88	1	0.53	3	1.59	0	0.00
Year 8	Boys	200	100	126	63.00	52	26.00	15	7.50	4	2.00	2	1.00	1	0.50
	Girls	172	100	108	62.79	43	25.00	12	6.98	4	2.33	1	0.58	4	2.33

† No answer.

in all of the groups that completed *Space Game* or *Question Time*.

In this section, I first look at how the children felt about their NEMP placement and how important boys and girls felt in the different group types. I then examine how much boys and girls enjoyed the two tasks in the different group types and how they felt about the workload in their respective groups. Finally, I explore how children in the different group types evaluated their group performance. For each question, I first present the overall results for the two gender groups at the two ages working on the two tasks. I then separately present the means for boys and girls in the different group types.

NEMP PLACEMENT

The children indicated how they felt about their group placement using the five-point faces rating scale depicted above. **Table 39** shows that the majority of the children, both boys and girls, rated their group placement favourably (ratings 4–5) for both tasks at both ages. However, a higher proportion of Year 4 children used the highest rating.

The analysis by gender and group type in **Table 40** shows two things in particular. First, both boys and girls at both age levels liked their group placement the least in the group type where they were outnum-

TABLE 40 THE EXTENT TO WHICH CHILDREN LIKED BEING IN THEIR NEMP GROUP, BY GENDER AND GROUP TYPE

		OVERALL $\bar{X}$	4B $\bar{X}$	3B1G $\bar{X}$	2B2G $\bar{X}$	1B3G $\bar{X}$	4G $\bar{X}$
QUESTION TIME GROUPS							
Year 4	Boys	4.49	4.42^{1}	4.63^{a1}	4.53^{1}	3.96^{b1}	•
	Girls	4.59	•	4.08^{a2}	4.71^{2}	4.61^{b2}	4.75^{2}
Year 8	Boys	4.41	4.55^{1}	4.57^{1}	4.33^{1}	3.58^{1}	•
	Girls	4.54	•	4.26	4.54	4.64	4.59
SPACE GAME GROUPS							
Year 4	Boys	4.57	4.44^{2}	4.70^{2}	4.65^{2}	3.85^{2}	•
	Girls	4.65	•	4.43^{2}	4.65^{2}	4.76^{2}	4.50^{2}
Year 8	Boys	4.49	4.71^{1}	4.54^{1}	4.41^{1}	4.20^{1}	•
	Girls	4.51	•	4.31	4.57	4.55	4.38

Gender differences within group types::
[a] = $p < 0.05$;
[b] = $p < 0.01$.

Within-gender differences in the different group types
[1] = $p < 0.05$;
[2] = $p < 0.01$.

bered. This reinforces the idea that the children did not like being in the minority. Secondly, there was no clear indication that the children gained most enjoyment in the same-gender groups. The highest mean in the same-gender groups occurred only at Year 8, and then only for one of the tasks—*Question Time* for girls and *Space Game* for boys. The differences between the means in the different group types were significant for boys on both tasks at both age levels, but for girls only in the Year 4 groups.

TABLE 41 CHILDREN'S FEELINGS OF IMPORTANCE IN THEIR GROUPS ACCORDING TO STARS RATING SCALE

		N	%	***	%	**	%	*	%	NA†	%
QUESTION TIME GROUPS											
Year 4	Boys	208	100	104	50.00	77	37.02	27	12.98	0	0.00
	Girls	192	100	98	51.04	65	33.85	28	14.58	1	0.52
Year 8	Boys	193	100	63	32.64	115	59.59	14	7.25	1	0.52
	Girls	163	100	39	23.93	111	68.10	12	7.36	1	0.61
SPACE GAME GROUPS											
Year 4	Boys	219	100	139	63.47	54	24.66	26	11.87	0	0.00
	Girls	189	100	89	47.09	80	42.33	20	10.58	0	0.00
Year 8	Boys	200	100	49	24.50	133	66.50	15	7.50	3	1.50
	Girls	172	100	39	22.67	113	65.70	19	11.05	1	0.58

† No answer.

The children also indicated how important they felt in their group using a stars rating scale, where three stars meant very important and one star meant of little importance. **Table 41** shows that in both the *Question Time* groups and the *Space Game* groups, the majority of the children, both boys and girls, felt very important at Year 4 and moderately important at Year 8.

The analysis by gender and group type (**Table 42**) showed different patterns for boys and girls. At both ages in the *Question Time* groups, and at Year 8 in the *Space Game* groups, boys in the 1b3g groups felt less important than did boys in the other group types. Therefore, in three of the four cases, boys felt least important in the same group type where they were least happy about their placement. Girls in the same-gender groups felt least important in three of the four cases (at both ages in the *Question Time* groups and at Year 4 in the *Space Game* groups). However, the differences between the means in the different group types were not statistically significant.

TABLE 42 CHILDREN'S FEELINGS OF IMPORTANCE IN THEIR GROUPS, BY GENDER AND GROUP TYPE

		OVERALL *X*	4B *X*	3B1G *X*	2B2G *X*	1B3G *X*	4G *X*
QUESTION TIME GROUPS							
Year 4	Boys	2.37	2.46	2.36	2.41	2.17	•
	Girls	2.37	•	2.42	2.34	2.42	2.00
Year 8	Boys	2.26	2.36[a]	2.32	2.18	2.00	•
	Girls	2.17	•	2.20	2.09	2.36	2.07[a]
SPACE GAME GROUPS							
Year 4	Boys	2.52	2.46[a]	2.49	2.58	2.43	•
	Girls	2.37	•	2.29	2.42	2.41	1.92[a]
Year 8	Boys	2.17	2.32	2.17[b]	2.19	1.90[a]	•
	Girls	2.12	•	1.81[b][1]	2.14[1]	2.20[a][1]	2.19[1]

Gender differences within group types:
[a] = $p < 0.05$;
[b] = $p < 0.01$.

Within-gender differences in the different group types:
[1] = $p < 0.04$.

TABLE 43 CHILDREN'S ENJOYMENT OF THEIR GROUP TASK ACCORDING TO 5-POINT FACES RATING SCALE

		N	%	☺	%	😐	%	😐	%	😞	%	☹	%	NA[1]	%
QUESTION TIME GROUPS — QUESTION 1															
Year 4	Boys	208	100	126	60.58	42	20.19	18	8.65	6	2.88	15	7.21	1	0.48
	Girls	192	100	117	60.94	52	27.08	14	7.29	4	2.08	4	2.08	1	0.52
Year 8	Boys	193	100	74	38.34	77	39.90	30	15.54	9	4.66	3	1.55	0	0.00
	Girls	163	100	59	36.20	66	40.49	35	21.47	2	1.23	1	0.61	0	0.00
SPACE GAME GROUPS															
Year 4	Boys	219	100	162	73.97	36	16.44	10	4.57	4	1.83	4	1.83	3	1.37
	Girls	189	100	132	69.84	41	21.69	10	5.29	1	0.53	4	2.12	1	0.53
Year 8	Boys	200	100	104	52.00	59	29.50	22	11.00	10	5.00	4	2.00	1	0.50
	Girls	172	100	66	38.37	75	43.60	24	13.95	6	3.49	0	0.00	1	0.58

ENJOYMENT OF THE GROUP TASK

To gauge whether the children enjoyed their group task equally in the different group types, I asked them to indicate their enjoyment of the task using the five faces rating scale. **Table 43** shows that, at both ages, the majority of the children, both boys and girls, rated the task they worked on favourably (ratings 4–5). In both the *Question Time* and the *Space Game* groups, however, a higher proportion of Year 4 than Year 8 children gave the task the highest rating. It is also worth noting that the proportions of boys and girls in the different categories were very similar in all cases.

The analysis by gender and group type (**Table 44**) shows that the boys in the same-gender groups at Year 4 enjoyed both tasks the least, whereas the boys in the same-gender groups at Year 8 enjoyed the tasks the most. Among the mixed-gender groups, boys in the gender-balanced groups enjoyed the task the most and boys in the 1b3g groups enjoyed the task the least in three of the four cases (at both ages in the *Question Time* groups and at Year 4 in the *Space Game* groups). The differences between the means for boys in the different group types were statistically significant for both tasks at Year 4 and in the *Space Game* groups at Year 8.

TABLE 44 CHILDREN'S ENJOYMENT OF THEIR GROUP TASK, BY GENDER AND GROUP TYPE

		OVERALL $\bar{X}$	4B $\bar{X}$	3B1G $\bar{X}$	2B2G $\bar{X}$	1B3G $\bar{X}$	4G $\bar{X}$
QUESTION TIME GROUPS							
Year 4	Boys	4.25	3.96[1]	4.28[1]	4.37[1]	4.00[1]	•
	Girls	4.43	•	4.54	4.33	4.49	4.75
Year 8	Boys	4.09	4.32	4.02	4.08	3.69[a]	•
	Girls	4.10	•	3.80	4.04	4.31[a]	4.21
SPACE GAME GROUPS							
Year 4	Boys	4.61[2]	4.48[2]	4.58[2]	4.72[2]	4.48[2]	•
	Girls	4.57	•	4.48[1]	4.60[1]	4.56[1]	4.67[1]
Year 8	Boys	4.25	4.61[2]	4.38[2]	4.00[2]	4.05[2]	•
	Girls	4.18	•	4.23	4.23	4.08	4.19

Gender differences within group types:
[a] = $p < 0.05$.

Within gender differences in the different group types:
[1] = $p < 0.05$;
[2] = $p < 0.01$.

Girls at Year 4 enjoyed both tasks the most in the same-gender groups. When they were outnumbered (in the 3b1g groups), they enjoyed the task the least in two of the four cases (in the *Question Time* groups at Year 8 and in the *Space Game* groups at Year 4). The differences between the means for girls in the different group types were statistically significant only for the *Space Game* groups at Year 4.

TABLE 45 CHILDREN'S RATINGS OF FAIRNESS IN THEIR GROUPS

		N	%	VERY FAIR	%	QUITE FAIR	%	NOT FAIR	%	NA†	%
QUESTION TIME GROUPS											
Year 4	Boys	208	100	116	55.77	77	37.02	14	6.73	1	0.48
	Girls	192	100	122	63.54	56	29.17	14	7.29	0	0.00
Year 8	Boys	193	100	111	57.51	69	35.75	9	4.66	4	2.07
	Girls	163	100	77	47.24	80	49.08	5	3.07	1	0.61
SPACE GAME GROUPS											
Year 4	Boys	219	100	148	67.58	47	21.46	21	9.59	3	1.37
	Girls	189	100	129	68.25	55	29.10	5	2.65	0	0.00
Year 8	Boys	200	100	122	61.00	74	37.00	4	2.00	0	0.00
	Girls	172	100	106	61.63	61	35.47	1	0.58	4	2.33

†No answer.

Although this analysis indicated that the group types where the tasks were enjoyed most were also the group types where children were happiest about their placement, this relationship was not very consistent.

THE WORKLOAD IN DIFFERENT GROUP TYPES

The children evaluated the workload in their group by indicating how fair they felt things were in their group and how happy they were with their workload. The children gave an indication of fairness in their groups using a three-point rating scale: very fair, quite fair and not fair. **Table 45** shows that in the *Question Time* groups at Year 4 and in the *Space Game* groups at both ages, the majority of the children, both boys and girls, said that they felt that things were very fair in their group. A small proportion of children, mostly from Year 4, felt that things were not fair in their group.

TABLE 46 CHILDREN'S RATINGS OF FAIRNESS IN THEIR GROUPS, BY GROUP TYPE

		OVERALL $\bar{X}$	4B $\bar{X}$	3B1G $\bar{X}$	2B2G $\bar{X}$	1B3G $\bar{X}$	4G $\bar{X}$
QUESTION TIME GROUPS							
Year 4	Boys	2.49	2.42	2.41	2.57	2.54	•
	Girls	2.56	•	2.29	2.69	2.56	2.00
Year 8	Boys	2.54	2.57	2.47	2.58	2.54	•
	Girls	2.44	•	2.45	2.51	2.38	2.36
SPACE TIME GROUPS							
Year 4	Boys	2.59	2.64	2.61	2.55	2.57	•
	Girls	2.66	•	2.50	2.66	2.70	2.75
Year 8	Boys	2.59	2.53	2.69	2.53	2.50	•
	Girls	2.63	•	2.52	2.70	2.64	2.44

As **Table 46** shows, the boys' mean ratings were very similar in the different group types at the two ages for both the *Question Time* and the *Space Game* groups. Therefore, boys in the different group types felt that things were equally fair. Except for the *Space Game* groups in Year 4, girls were least positive about fairness in the same-gender groups and most positive about fairness in the balanced mixed-gender groups.

TABLE 47 CHILDREN'S FEELINGS ABOUT THEIR SHARE OF THE GROUP WORKLOAD, USING THE 5-POINT FACES RATING SCALE

		N	%	😊	%	🙂	%	😐	%	🙁	%	☹	%	NA†	%
QUESTION TIME GROUPS															
Year 4	Boys	208	100	140	67.31	44	21.15	13	6.25	5	2.40	5	2.40	1	0.48
	Girls	192	100	136	70.83	40	20.83	8	4.17	3	1.56	4	2.08	1	0.52
Year 8	Boys	193	100	110	56.99	61	31.61	18	9.33	2	1.04	1	0.52	1	0.52
	Girls	163	100	98	60.12	52	31.90	10	6.13	1	0.61	1	0.61	1	0.61
SPACE GAME GROUPS															
Year 4	Boys	219	100	154	70.32	43	19.63	10	4.57	6	2.74	5	2.28	1	0.46
	Girls	189	100	147	77.78	30	15.87	4	2.12	2	1.06	5	2.65	1	0.53
Year 8	Boys	200	100	119	59.50	60	30.00	19	9.50	1	0.50	–	0.00	1	0.50
	Girls	172	100	111	64.53	46	26.74	11	6.40	3	1.74	–	0.00	1	0.58

†No answer.

In regard to how happy the children felt with their share of the workload, **Table 47** shows that, in both tasks, the majority of the children, both boys and girls, were happy with their share of the work (ratings 4–5). In both the *Question Time* and the *Space Game* groups, a larger proportion of Year 4 than Year 8 children used the highest rating. Once again, the proportions of boys and girls in the different categories were very similar in all cases.

In all cases, boys in the same-gender groups felt happiest about their share of the work (**Table 48**). However, when girls outnumbered boys in the mixed-gender groups, the boys consistently felt less happy than they did in the other group types. The differences between the means for boys in the different group types were statistically significant in the *Space Game* groups at both ages. Girls in the same-gender groups felt least happy about their share of the work in three of the four cases (at Year 8 in the *Question Time* groups and at both ages in the *Space Game* groups). The means for girls were very similar in the three mixed- gender group types in all four cases.

Essentially, this analysis showed that while boys were happiest with their share of the workload in the same-gender groups, girls were least happy with their share in the same-gender groups.

TABLE 48 CHILDREN'S FEELINGS ABOUT THEIR SHARE OF THE GROUP WORKLOAD, BY GENDER AND GROUP TYPE

		OVERALL $\bar{X}$	4B $\bar{X}$	3B1G $\bar{X}$	2B2G $\bar{X}$	1B3G $\bar{X}$	4G $\bar{X}$
QUESTION TIME GROUPS							
Year 4	Boys	4.49	4.67^{2}	4.34^{2}	4.63^{2}	4.29^{2}	•
	Girls	4.58	•	4.46^{1}	4.64^{1}	4.50^{1}	4.88^{1}
Year 8	Boys	4.44	4.57^{2}	4.45^{2}	4.45^{2}	3.92^{b2}	•
	Girls	4.51	•	4.55	4.47	4.67^{b}	4.37
SPACE TIME GROUPS							
Year 4	Boys	4.54	4.57	4.57	4.56	4.24	•
	Girls	4.66	•	4.75	4.65	4.66	4.50
Year 8	Boys	4.49	4.68^{a}	4.53	4.41	4.35	•
	Girls	4.55	•	4.65	4.58	4.55	4.25^{a}

Gender differences within group types:
^{a} = $p < 0.05$;
^{b} = $p < 0.01$.

Within-gender differences in the different group types:
^{1} = $p < 0.05$;
^{2} = $p < 0.01$.

TABLE 49 CHILDREN'S RATINGS OF TASK PERFORMANCE, USING THE FIVE-POINT FACES RATING SCALE

		N	%	☺	%	😐	%	😐	%	😐	%	☹	%	NA†	%
QUESTION TIME GROUPS															
Year 4	Boys	208	100	105	50.48	71	34.13	18	8.65	7	3.37	6	2.88	1	0.48
	Girls	192	100	106	55.21	65	33.85	13	6.77	6	3.12	1	0.52	1	0.52
Year 8	Boys	193	100	82	42.49	75	38.86	27	13.99	5	2.59	3	1.55	1	0.52
	Girls	163	100	79	48.47	65	39.88	15	9.20	4	2.45	–	0.00	–	0.00
SPACE GAME GROUPS															
Year 4	Boys	219	100	120	54.79	66	30.14	17	7.76	6	2.74	8	3.65	2	0.91
	Girls	189	100	103	54.50	71	37.57	7	3.70	4	2.12	2	1.06	2	1.06
Year 8	Boys	200	100	89	44.50	90	45.00	17	8.50	4	2.00	–	0.00	–	0.00
	Girls	172	100	89	51.74	70	40.70	10	5.81	2	1.16	–	0.00	1	0.58

†No answer.

TASK PERFORMANCE IN DIFFERENT GROUP TYPES

Finally, the children evaluated how well they thought their group did on the particular task. Once again, they responded using the five faces rating scale. At both ages (see **Table 49**), the majority of boys and girls thought their group did well on their task (ratings 4–5), and the proportions of boys and girls in the different categories were very similar in all cases.

The analysis by gender and group type (**Table 50**) revealed that in both the *Question Time* and *Space Game* groups, the boys in the same-gender groups had their lowest mean at Year 4 and their highest mean at Year 8. There was no particular pattern for boys in the mixed-gender groups. Statistically significant differences between the means for boys for the different group types were evident for the *Question Time* groups at both ages. Girls in the same-gender groups had their highest mean at Year 4, but no pattern was observed at Year 8. Among the mixed-gender groups, girls in the 3b1g groups consistently felt they did least well. The difference between the means for the girls in the different groups were statistically significant for the *Question Time* groups at Year 8 and the *Space Game* groups at Year 4.

TABLE 50 CHILDREN'S RATINGS OF TASK PERFORMANCE, BY GENDER AND GROUP TYPE

		OVERALL $\bar{X}$	4B $\bar{X}$	3B1G $\bar{X}$	2B2G $\bar{X}$	1B3G $\bar{X}$	4G $\bar{X}$
QUESTION TIME GROUPS							
Year 4	Boys	4.27	3.96[1]	4.10[1]	4.44[1]	4.46[1]	•
	Girls	4.41	•	4.08	4.48	4.42	4.50
Year 8	Boys	4.19	4.40[1]	4.30[1]	4.00[1]	4.08[a1]	•
	Girls	4.34	•	4.15[1]	4.22[1]	4.56[a1]	4.50[1]
SPACE TIME GROUPS							
Year 4	Boys	4.31	4.04[a]	4.24	4.47	4.30	•
	Girls	4.44	•	4.26[1]	4.39[1]	4.56[1]	4.58[a1]
Year 8	Boys	4.32	4.47	4.31	4.29[a]	4.25	•
	Girls	4.44	•	4.32	4.54[a]	4.42	4.31

Gender differences within group types:
[a] = $p < 0.05$.

Within gender differences in the different group types:
[1] = $p < 0.01$.

In general, this analysis showed that boys and girls as sub-groups did not evaluate their group experience in the various groups similarly. Moreover, the children's evaluation of their performance in a particular group generally did not extend to both tasks at both age levels.

SUMMARY

The children's evaluations of their group experiences highlighted several points of interest:

- Although the children's responses to the questions asked of them were generally positive, the analysis by group type showed that children in different groups sometimes felt differently about a particular point. For example, although the majority of the children rated their placement highly, both boys and girls expressed least positive views about being in the group type where they were outnumbered.
- Certain trends related to the children's age level more than to their group type. For example, the Year 4 children liked the activities more than did the Year 8 children.
- For boys, the group experience was least positive in the 1b3g groups. For girls, however, the group experience was not more positive in the same-gender groups than in the mixed-gender groups. In fact, girls in the 4g groups felt least important, and they also felt that things were least fair in that group type.

Overall, the children's responses and attitudes were neither consistently positive nor negative about particular group types. However, the Year 8 children were more positive than the Year 4 children about working in balanced mixed-gender groups. Boys generally enjoyed working in the same-gender groups more than did girls. Girls generally were more positive about working in mixed-gender groups as long as they were not outnumbered.

At both age groups, boys and girls were least positive about working in groups where they were in the minority. This attitude showed up in the children's general views and in their post-task evaluations. This negative attitude persisted even after the NEMP group experience, although it is important to point out that the disadvantaged position of the outnumbered student was frequently perceived rather than real. Moreover, boys and girls did not appear to be very enthusiastic in general about working in balanced mixed-gender groups, especially at Year 4. However, their post-task evaluations indicated that they rated their experience in these groups as relatively positive. Hence, there seemed to be a change in attitude *after* the experience. This latter finding indicates the importance and value of giving young children the opportunity to work in same-gender and balanced mixed-gender groups.

7

DISCUSSION AND CONCLUSIONS

This concluding chapter compares the findings of the current study with previous research findings. Possible implications for several educational theories relating to gender and group work are considered, and the limitations of this study are outlined. The report ends with several conclusions on the ways in which this work has advanced our knowledge about studying group work and discusses several implications for structuring this activity with children in educational settings.

PARTICIPATION

In the current stu dy, the results revealed no clear differentiation (at both Years 4 and 8) in terms of the children's participation between the various same-gender and mixed-gender groups. The level of participation was not consistently higher in the same-gender groups than in the other groups, a finding at odds with the results reported by Jacklin and Maccoby (1978) and McCloskey and Coleman (1992) in their studies on play-groups. Nor was participation consistently higher in the balanced mixed-gender groups than in the same-gender groups, a finding at variance with that of Aries (1982) in her study of adult discussion groups in a face-to-face situation and of Savicki, Kelley and Lingenfelter (1996a) in their computer-mediated-communication (CMC) setting. Moreover, the present study revealed that participation in the three mixed-gender group types was neither consistently higher nor consistently lower than in the same-gender groups. However, the children's age did influence their participation in particular group types, with the members of the 4g groups participating more in the older age group than in the younger age group, and the members of the 2b2g groups doing the opposite.

The results of the analysis of the mean participation levels of boys and girls in the different group types indicated that, at Year 4, boys and girls participated less in the group type where they were outnumbered than in the other group types. These results support Johnson and Shulman's (1989) conclusion that the participation of females and males decreases in the group type where they are outnumbered, and that their participation level is lower in that group type than in the other group types. At Year 8, both gender groups, the girls especially, participated less in the 2b2g groups than in the other group types. This result differs from those studies where boys and girls were found to participate less in mixed-gender groups in general (e.g., Jacklin and Maccoby, 1978; McCloskey and Coleman, 1992).

Although the results for boys across the four group types revealed no consistent differentiation in levels of participation between the same-gender and the mixed-gender groups, they did indicate that the boys' participation level tended to be lower in one of the group types than in the others. In neither age group did the boys' results consistently support Carli's (1989) observation that males tend to participate more in same-gender groups than in mixed-gender groups. Nor did they support Aries' (1982) converse observation of higher male participation in mixed-gender than in same-gender groups.

The results for girls showed no consistent differentiation in participation levels across the two age groups. At Year 4, girls' levels of participation were greater in the mixed-gender groups where they were not outnumbered than in the same-gender groups. This finding supports the results of Carli's (1989) study with college discussion groups and Kutnick's (1997) study with children (ages 9–10). Taken together, these results do not support Webb's (1984) and Maccoby's (1998) proposition that girls are likely to participate more in same-gender groups. Maccoby, in fact, argues that girls' participation is likely to be favoured in a same-gender group because 'a reduction of contact between the two sexes during middle childhood ... protects girls from male domination and coercion' (p73). At Year 8, however, girls' highest levels of participation occurred in the same-gender groups. This result accords with Maccoby's (1998) developmental view that, at certain stages of their lives, children tend to exhibit a higher preference for their own gender than for the opposite gender and that this leads to higher participation in same-gender groups.

The overall results of the analysis of boys' and girls' participation levels in the different group types support Webb's (1984) assertion that disparity between the participation of males and females is smaller in balanced mixed-gender groups than in imbalanced mixed-gender groups. However, in the current study, there was no evidence of domination by males in the mixed-gender groups across the three tasks. More particularly, there was no general indication, at either age level, of females' contributions in children's mixed-gender task groups being interrupted, overlooked, ignored or unheard, as suggested by Butler and Geis (1990). Rather, there appeared to be traces of a relationship between the subject area and the participation levels of boys and girls, with boys participating more than girls in the first discussion of the technology task and girls participating more than boys on the language task. Holden (1993) reached similar though stronger conclusions in her study involving mathematics/technology and language tasks.

One result from the analysis of the contribution of ideas by boys and girls in the different group types is worth discussing here. In *Space Game's* Discussion One, boys generated more ideas than girls in all three mixed-gender groups at Year 4. This result supports Strodtbeck and Mann's (1955) finding that males contribute more suggestions than females irrespective of the number of females in the group. The present study, however, acknowledges that the subject area and the content of the activity (improving a board game on space) may have contributed to this disparity in the generation of ideas by boys and girls. Also, this pattern was evident in only one of the two discussions that made up the technology task, and it was not repeated at Year 8 in the same task or in either of the other two tasks at both age levels.

The results regarding the children's ability to organise a task showed that in two of the three tasks (*Separating Mixtures* and

Question Time) more girls than boys were involved in organisation at both age levels, whereas in *Space Game* more boys than girls were involved in organisation at both age levels. These results suggest that involvement in organisation was related more to the nature of the task than to gender characteristics as suggested by writers such as Carli (1982) and Aries (1982). Moreover, because in this study organisation was coded differently from on-task participation, the results showed an interesting pattern. When students (boys and girls) became highly involved in organisation, their on-task participation tended to be low and vice-versa. Therefore, this finding suggests that in addition to monitoring the involvement or otherwise of students during group activities, teachers should also monitor the kind of on-task behaviour so as to ensure that students participate and develop different group skills.

INTERACTION

The interaction results of the current study were different for the Year 4 and the Year 8 groups, thus supporting Leaper's (1991) observation of inconsistencies between the interaction patterns of five- and seven-year-old children, even though the groups were involved in the same task. The Year 8 results, which revealed that the 2b2g groups had a lower interaction mean than the other group types, support the findings of other studies, such as those by Tolmie and Howe (1993) and Underwood, McCaffney and Underwood (1990). It should be noted, however, that these latter studies compared same-gender and balanced mixed-gender groups.

In my study, the same-gender groups generally had a higher interaction mean than the mixed-gender groups. (This finding has been reported by researchers carrying out studies in other contexts, such as Jacklin and Maccoby's (1978) study in a play context and Tolmie and Howe's (1993) and Underwood et al.'s (1990) studies carried out in IT settings.) However, when the results were separated out for the two activities in which interaction patterns were examined, a higher level of interaction in the same-gender than in the mixed-gender groups was found for only one activity at both age levels.

Overall, across the two activities in the two age groups, the 4b groups were observed to interact the most. This finding supports the result of the study by Tolmie and Howe (1993), but it does not support the results of several others studies, such as those by Jacklin and Maccoby (1978), Lee (1993) and Pryor (1995), which reported higher interaction in the girls' groups.

CO-OPERATION

In general, in the current study the older groups co-operated more than the younger ones across all the group types. These results do not support Tann's (1981) observation of reduced co-operation during English tasks in older mixed-gender groups. Furthermore, my study found that the amount of co-operation observed in the different activities that made up each of the tasks varied.

The 4g groups had the highest co-operation mean at both Years 4 and 8. This finding, which reflected a number of activities at both ages, supported the conclusions reached by researchers such as Wood (1987) and Underwood, Jindal and Underwood (1994). It also supported Maccoby's observation that girls' speech is in general more co-operative in nature. However, overall, there was no indication that the same-gender groups co-operated more than the mixed-gender groups, as reported by Underwood et al. (1994), or that they co-operated less, as predicted

by Hoffman (1965) and supported by studies such as that by Wiley (1973) cited in Meeker and Wietzel-O'Neil (1977).

CONFLICT

The overall results of conflict in the current study suggested that conflict was present almost evenly in all of the group types at both age levels. This analysis did not support Carli's (1989) finding that conflict is more of a phenomenon in mixed-gender group types. The Year 4 results indicating that the 4g groups had a higher overall conflict mean than the other group types support Quicke and Winter's (1995) observation that girls do have conflicts in same-gender groups and that they do not leave matters unresolved as had been suggested by writers such as Tolmie and Howe (1993), who argued that girls' interaction tends to be conflict-free.

My study also showed that conflict was more common in the Year 4 than in the Year 8 groups. This finding has not been reported on by existing studies, which normally have examined only one age group (see, for example, Wilkinson, Lindow & Chiang, 1985), or else have grouped the results of different age cohorts together (see, for example, Miller, Dahaner & Forbes, 1986).

The analysis of conflict also showed that the percentage of groups experiencing conflict differed across the various activities that made up each of the three tasks. In *Separating Mixtures* there was more conflict during the discussion than the experiment. In *Question Time* there was more conflict during the question choice activity than the reporting/justification activity, and in *Space Game* the most conflict occurred in the game, and there was more in Discussion One than in Discussion Two. Savicki, Kelly & Lingenfelter (1996b) similarly noted a possible relationship between the nature of the task and the amount of conflict. However, what my study suggests is that the amount of conflict varies even within activities that make up one task. It also suggests that the amount of conflict in particular group types differs with the nature of the activity. The 4g groups became highly involved in conflict during verbal activities, while the 4b groups became highly involved in conflict during activities where they were doing something physical (carrying out an experiment, playing a board game).

Finally, the results of the analysis of conflict support the conclusions reached by Wilkinson et al. (1985) that (i) there are no differences between boys and girls in terms of how much they initiate and participate in conflicts, and (ii) that the majority of the conflicts that take place in mixed-gender groups are mixed-gender conflicts. My observation that in many cases conflicts included more than two group members shows a limitation in Miller et al.'s (1986) study, which recorded all conflicts as involving two participants in groups with six members.

ACHIEVEMENT

At Year 4, the achievement of the same-gender groups was lower than the achievement of the mixed-gender groups. These scores do not support Underwood and his colleagues' (1990, 1994) conclusion that same- gender groups perform better than mixed-gender groups on group tasks. Rather, the results from the Year 4 groups support Wood's (1987) conclusion that mixed-gender groups tend to perform better than same- gender groups. However, this pattern was not evident for the Year 8 groups, where no difference was apparent between the same- and the mixed-gender groups.

The analysis of the relationship between the mean performance scores and the means for interaction, ideas, co-operation and conflict levels in the different group types did not yield consistent results across the three tasks and at the two age levels. It therefore was impossible to make any generalisations about the relationship between processes and products. The most common relationship recorded was the inverse relationship between group conflict and the performance mean. This finding does not support the assertion made by Johnson and Johnson (1994) that conflict renders the group experience more effective.

ATTITUDES

The questionnaire results of the current study showed that the children's attitudes towards same-gender and mixed-gender groups changed with age. At Year 4, both boys and girls most enjoyed working in same- gender groups. Although they enjoyed working in gender-balanced groups less, they preferred this group setting to the one where they worked with members of the other gender only. These results support Whiting and Edwards' (1988) finding of preference for same-gender groups. At Year 8, however, boys and girls responded equally favourably to same-gender and balanced mixed-gender groups, although they still felt less positive about working with members of the other gender only. The fact that boys and girls responded more favourably to working in balanced mixed-gender groups at Year 8 than at Year 4 does not support Whiting and Edwards' (1988) or Jacklin and Maccoby's (1987) observations that same-gender preference increases with age. However, it is worth noting that the former studies were carried out over a decade ago and, as Maccoby (1998) points out, the cultural context changes over time. The results of the present study therefore may indicate a general cultural change or else a pattern specific to New Zealand children.

At both ages, more girls than boys said that they enjoyed working in balanced mixed-gender groups. The difference was statistically significant for both age groups. Similar findings were reported in studies carried out by Barbieri and Light (1992) and Pryor (1995) in the United Kingdom, although the student samples in these studies were much smaller than the sample in the current study. The issue of girls being more willing to interact with boys than vice versa has also been suggested by Maccoby (1998).

At both ages, working in a group with members of the other gender only was the least attractive option for both boys and girls. According to Maccoby (1998), this finding relates to a developmental issue in school-age children. Maccoby claims that children of this age sense that there are things about the social world of the other gender group that are unknown or not understood. As such, they lack confidence about how to interact with members of the other gender, and they see no reason why they should want to interact with them. This leads to children becoming alienated from the other gender, an alienation that reaches its peak at about eight years of age.

THE POST-TASK EVALUATION IN THE DIFFERENT GROUP TYPES

In regard to the children's post-task evaluation of the group experience in the different group types, three main points are worth discussing. First, the children's responses indicated that the experience tended to be less positive for the boy or the girl in the minority in the 1b3g and 3b1g groups. These students were less positive about their group placement, their impor-

tance in the group and their enjoyment of the task. The boys in the 1b3g groups were also less happy about their share of work than were the boys in the other mixed-gender groups. Among the girls in the mixed-gender groups, the lowest rating for group performance came from the girls in the 3b1g group.

The second point is that in three of the four student samples, girls in the 4g groups compared to girls in the other group types felt least important, thought things were least fair and were least happy with their share of the work. The girls' responses did not suggest that girls in the 4g group type felt protected from male domination, as suggested by Maccoby (1998).

Third, although the children, especially at Year 4, said that same-gender groups were their most preferred group type, once they experienced working in balanced mixed-gender groups their evaluation of their experience became relatively positive. It appears that once the children experienced working with both male and female group members, their fear and/or discomfort of working with children of the other gender diminished. From the children's perspectives, the experiences in the same-gender and the balanced mixed-gender groups were the more enjoyable and the more productive. However, the children's perceptions of discomfort and insecurity about being outnumbered and working in a group with members of the other gender only remained strong even after the experience in the imbalanced mixed-gender groups. Overall, the nature of the tasks did not play an influential role in the children's evaluations, and similar results were achieved in the *Question Time* and the *Space Game* samples.

INTERPRETING THE FINDINGS USING VARIOUS EDUCATIONAL THEORIES

Although no one existing theory provides a convincing explanation for the findings of the present study, aspects of five different theories that discuss gender issues in group work have relevance here. These theories are expectation states theory, social role theory, structural numerical proportions theory, group cognition theory and post-structuralist theory.

EXPECTATION STATES THEORY

This theory is relevant to the current study because it is specifically concerned with task-oriented groups working together on collective tasks (Dion, 1985). This theory uses the social psychology of expectations and the concept of status characteristics to account for the interactional inequalities of power and influence produced by gender (Berger, Cohen & Zelditch, 1972; Berger, Rosenholt & Zelditch, 1980). Expectation states theory regards gender as socially constructed and suggests that its impact on the group depends on the particular situation in which the group finds itself. In general, this theory predicts greater variation between males and females in mixed-gender groups than between males and females in same-gender groups. In mixed-gender groups, females are expected to speak less, offer fewer suggestions and be less influential overall.

On several occasions the results of my study did not support these predictions. In the three tasks, at both age levels, the less involved group members in the mixed-gender groups sometimes were males and sometimes were females. Moreover, the difference between the mean participation levels of males and females was largest between the two same-gender groups in the science task at Year 4. In this case, especially, gender did not act as a status char-

acteristic because the differences between the participation means of males and females were smaller in the mixed-gender than in the same-gender groups. Moreover, in the 2b2g groups, in particular, males were not observed to be consistently more involved in the tasks than were females.

The results of my study did, however, support expectation states theory during one activity. In Discussion One of the technology task, boys were observed to act with a higher degree of agency than the girls in that they provided the greater number of ideas in all three mixed-gender groups at Year 4. Expectation states theory predicts that males will be more dominant than girls in mixed-gender groups and also that the nature of the activity within the group will influence the degree of dominance. What the present study adds is that higher agency by males in this case resulted from the *content* of the discussion rather than the *nature* of the activity, for there were other discussions where similar participation patterns did not occur.

SOCIAL ROLE THEORY

Social role theory (see Eagly, 1987) is included primarily because it offers different predictions from those of expectation states theory about performance in same-gender groups. Essentially, social role theory, which was developed in the 1980s, sees society as being gendered, with males and females having different gender cultures and fulfilling different roles in society. This theory predicts that males will react with a higher degree of agency and females will react with a more communal spirit in groups. In same-gender groups, the prediction is that differences will be apparent between the male and the female groups. In mixed-gender groups, males are expected to participate more in task-related behaviour than females.

On several occasions, the results of this study did not support higher male agency in mixed-gender groups. Furthermore, in the three tasks at both age levels, the less involved group members in the mixed-gender groups were both males and females. In the 2b2g groups, in particular, males were not observed to be consistently more involved in the tasks than females, and during the board game in the technology task, males and females did not differ in giving help related to procedures to peers across groups with different gender compositions. Therefore, in this study, help was not observed to be part of the communal gender culture of females, as suggested by social role theory.

At the same time, two aspects of the results of this study do support social role theory. First, in Discussion One of the technology task, boys provided more ideas than girls in all of the three mixed-gender group types at Year 4. However, this pattern only occurred in one discussion in one task and was not repeated in the other discussion of the same task or in either of the two other tasks. Second, the difference between the mean participation levels of males and females was larger in the two same-gender groups than in the mixed- gender groups during the science task at Year 4.

My study also found that differences such as these relate more to the age of the students than to specific gender characteristics or to the nature of the task. At Year 4, the 4g groups had the lowest participation mean in two of the three tasks (in the discussion of the science task and in the brainstorming and the question choice activities of the language task). Then, at Year 8, the 4g groups had the highest mean in two of the three tasks (the discussion and the experiment of the science task and in both discussions of the technology task). An opposite pattern was apparent in the 4b

groups, although it was not as strong. At Year 4, the 4b groups had the highest mean in both activities of the science task, whereas at Year 8 they had the lowest mean in both discussions of the technology task.

STRUCTURAL NUMERICAL PROPORTIONS THEORY

This theory makes an important contribution to the study of group work for it does not classify all mixed- gender group types together. Kanter (1977) argues that groups consisting of varying proportions of males and females produce certain interaction patterns that tend to disadvantage those who are numerically in the minority. This theory predicts that the minority student will be disadvantaged in imbalanced mixed-gender groups, and that more equitable participation will take place in the balanced mixed-gender groups. This theory was developed mainly to point out the disadvantage experienced by minority females in male majority groups.

The present study shows a limitation of this theory, in that inequitable participation was found to be a characteristic of all group types, with both males and females 'sidelined' in the different mixed-gender groups. However, as predicted by structural numerical proportions theory, my study also found that there was a tendency at Year 4, especially in the 1b3g groups, for the outnumbered student to be involved, on average, to a lower degree than the other members in the gender-imbalanced groups, and/or to have the lowest mean participation level of that gender group when the means of the gender group were compared across group types.

GROUP COGNITION THEORY

Presented in Maccoby's (1998) book *The Two Sexes: growing up apart, coming together*, group cognition theory is relevant to my study because it attributes a powerful socialising role to peers. Also, Maccoby's views on children's groups are supported by very little task group research—an area that my study covers.

The traditional view of socialisation holds that children are socialised through processes in which adults pass on to each generation of children the rules, values and beliefs governing social behaviour in their culture. However, according to Maccoby (1998, 9), 'the socialisation account has not proved adequate to the task of explaining gender differentiation. The socialisation account is not wrong—just too narrow, too limited.' She suggests that this understanding brings about a shift of emphasis

> from the individual to the dyad or the larger social group. Sex linked behaviour turns out to be a pervasive function of the social context in which it occurs ...It turns out that the relevant condition is the gender composition of the social pair or group within which the individual is functioning at any given time. The gendered aspect of an individual's behaviour is brought into play by the gender of others. (p9)

Maccoby sees gender differentiation as a developmental issue, with boys and girls diverging at certain stages of development and converging at others. The preference for same-gender playmates is found as early as age three (Howes & Phillipsen, 1992; Pitcher & Schultz, 1983) and continues among school-age children, especially in contexts not controlled by adults. As mentioned above, during this period of their life, children lack the confidence to talk to or interact with someone of the other gender group, and they do not see any reason why they should want to do this. As Maccoby (1998, 62) observes, 'This sense of alienation from the other sex is an outcome of the gender segregation that has been a fact of life for most children during the years preceding adolescence.'

Group cognition theory predicts higher participation levels in same-gender than in mixed-gender groups. It also suggests that interaction in girls' groups is more co-operative than in boys' groups and that girls are better off in same-gender groups where they are protected from male domination. The experience for boys, however, is not expected to be different in same- and mixed-gender groups. This theory also suggests that girls are more willing to interact with boys than boys are with girls, and that boys and girls can work comfortably together in situations structured by adults because their student role becomes salient in such situations. However, it also suggests that the discourse of boys and girls is different in situations of dominance. Therefore, it is not clear whether domination is an issue for certain group types or for all group types in situations where group work activity is structured by adults but the children are left to work on their own for most of the time.

In line with this theory, boys and girls in the older groups in my study did participate more and accomplish higher levels of achievement on the tasks in the same-gender groups. Girls also were observed to be more co-operative and more willing to interact with boys. Furthermore, the minority boys and girls in the gender-imbalanced groups did express feelings of discomfort and insecurity about working with members of the opposite gender.

At variance with Maccoby's claim that school-age children continue to prefer same-gender groups was my finding that same-gender preference did not increase with age, and that Year 8 students were more willing than Year 4 students to be members of gender-balanced groups. Moreover, girls' experiences in the 4g groups were not necessarily more productive and/or more enjoyable than girls' experiences in the mixed-gender groups. The video analysis showed that, especially at Year 4, the 4g groups had the lowest participation means during several activities. The Year 4 girls participated more in the mixed-gender groups where they were not outnumbered than in the same-gender groups. Also, the post-task evaluation showed that girls in same-gender groups generally felt less important than girls in mixed-gender groups, felt that things were least fair and expressed the highest degree of unhappiness with their share of the work. Taken together, these results suggest that girls in the 4g groups neither evaluated their experiences more positively than girls in the other groups nor participated in and achieved more in this group type.

Furthermore, my study did not find the experience to be similar for boys in the different group types. In several activities, boys' levels of participation and interaction with other group members were greater in the same-gender groups. My study also shows that boys and girls can and do work comfortably together, especially in the 2b2g groups, and that domination is not an issue of concern during task-focused group work.

POST-STRUCTURALIST THEORY

I include post-structuralist theory primarily because of its recognition of the shifting, fragmented, multi-faceted and contradictory nature of human experiences and its move away from the view of unitary, non-contradictory selves (Davies, 1989). The lack of conclusiveness in the literature, my increasing awareness of the complexity of group work, and the different variables that seem to be at play during this activity encouraged me to explore ideas put forth by post-structuralism. According to Weiler (1988), post-structuralist feminist researchers increasingly recognise that students' identities cannot be reduced to one par-

ticular structural factor. Walkerdine (1981, 14) explains this position as follows:

> female teachers and small girls ... are no unitary subjects uniquely positioned but produced as a nexus of subjectivities, in relations of power which are constantly shifting, rendering them at one moment powerful and at the other powerless.

However, this viewpoint does not restrain Davies from placing gender as a more determining category than class and race.

From this perspective, people make their own sense from possible ways of being and out of the multitude of conflicting and often contradictory possibilities offered by society. Subjectivity is seen as a constantly changing process, and individuals therefore are allowed or even expected to behave differently in a similar context. At times in the present study, results varied within and across tasks and within and across age groups, with certain variables such as gender and/or the gender composition of the group sometimes becoming salient. At other times, however, there were no apparent trends and/or relationships between variables. Post-structuralist theory provides an explanation for these results through its suggestion that human behaviour is contradictory and fragmented and that gender does not emerge as a salient variable for all individuals (and groups) in a similar situation or for the same individuals at different times. The lack of consistency between the group achievement scores and the various group processes in my study certainly supports the notion of non-generalisation of behaviour that is posited by post-structuralist theory.

SUMMATION REGARDING FIVE THEORIES

My study throws light on the merit of five theories that discuss the impact of student gender in group work. Three of these theories (expectation states, social role and structural numerical proportions) when applied to the findings of this study appear to offer only generalities and do not support the findings consistently. While post-structuralist theory has relevance to my findings, it does not suggest any systematic patterns of relationships among variables, and so provides no guidance to practitioners about structuring group work. At present, an adequate theory to explain group work in classroom settings does not exist, although Maccoby's group cognition theory has substantial relevance for that context. Several of Maccoby's predictions are supported by the findings of this study (e.g., higher achievement in same-gender groups as children grow older; girls being more willing to interact with boys than vice versa; and feelings of discomfort and insecurity when in the minority). At the same time, however, the present study questions aspects of Maccoby's position within a learning context (e.g., the idea that the structure of boys' and girls' groups is different, as is the nature of their interaction in groups). It also indicates that there is a need to augment the types of data she has relied on with data relating to a variety of tasks in contemporary contexts. Even when updating and developing her theory (e.g., Maccoby, 1998), she still uses data that are up to 20 years old, and much of the research she cites is based in play settings.

LIMITATIONS OF THIS STUDY

The major limitation of my study is that it focused on the gender of the students in the different group types without incorporating the students' ethnicity and the socio-economic status of their schools into the analyses. The students' ethnicity is particularly important in a New Zealand context where there is concern about the diverse educational involvement of Pakeha, Maori, Asian and Pacific Island children. Although ethnicity and socio-economic data were available in the NEMP databases, attempts

at adding these variables to the overall study resulted in very small cell sizes and complex results. Consequently, I decided to focus solely on gender, given that the aim of my study was to present profiles of behaviour in the different group types across tasks and age levels. Data on the children's ethnicity and the socio-economic backgrounds of their schools will be used later in focused studies examining particular aspects of the results.

I am also aware that I have not provided the reader with detailed excerpts from raw data in my reporting on the different processes, and that I have looked at the interrelationships of the different processes (and products) in a very limited way. However, I had to be very selective in both the analyses and the write-up of this study. More work on the qualitative data available will be carried out in the future.

Also, it is important to keep in mind that this study was carried out in a performance assessment context, and that its findings are therefore most relevant to that context. As such, caution must be taken in interpreting the relevance of these results within the context of the normal classroom setting. That said, I would argue that the results of this study are more relevant to the classroom context than are the findings of laboratory conditions, and that they may be even more relevant than the findings of studies carried out in a play context.

Finally, after this research experience, I consider that studies of a similar nature involving the interpretation and analyses of interrelated group processes may benefit from having more than one researcher involved in these processes. The major advantage would be that of having someone with whom to construct meanings, clarify issues, set boundaries and discuss matters during the analysis process.

SUMMARY

The analysis of the three tasks, *Separating Mixtures, Question Time* and *Space Game,* showed that group dynamics did not remain consistent across the different activities that made up any one task. In order to present an accurate picture of the processes that occur at the different stages, each activity needs to be analysed separately. This observation highlights a limitation in previous research, where a very small amount of observation took place but the results were extrapolated to the period of the whole task, or where observation took place during a particular stage of the task but the results were assumed to represent the processes of the complete task. In my study, the group dynamics were not consistent across tasks in the different group types and therefore generalisations were not possible.

The video analysis showed that the group experience was not consistently more productive in the same- gender groups when compared to the balanced mixed-gender groups, nor was the experience equitable and productive in the gender-balanced groups when compared to the gender-imbalanced ones. Overall, no group type was observed to be more productive than any other group type.

The group experiences in 4g groups, in particular, provided results that are relevant to the debate regarding the merits of single-sex versus co-educational schooling. The results of this study do not support Mael's (1998) assertion that same-gender groupings benefit the academic achievement of girls. Girls in the 4g groups did not consistently achieve more than the other groups in terms of the group product. Even when achievement was extended to include participation on the tasks, girls in the 4g groups did not consistently participate more than did girls in the other group

types, especially at Year 4. Interestingly, the girls themselves did not evaluate the experience as being more positive in the 4g groups.

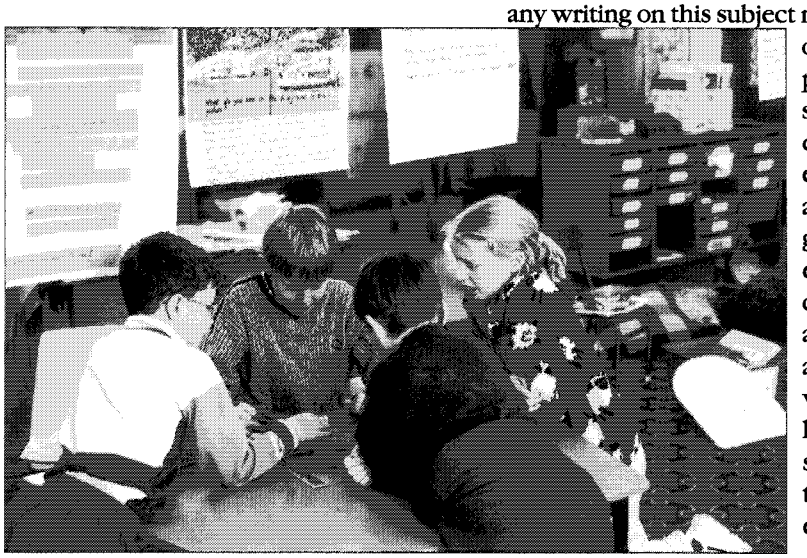
Mael (1998) also suggests that co-educational classes foster inequity and that same-sex classes would alleviate this problem. Similarly, Bailey (1993) sees same-sex schooling as the solution to classroom inequity. The present study shows that inequity occurs in both mixed-gender and same-gender groups and that same-gender groups do not provide students with more equitable participation patterns.

Overall, the analyses of this study revealed an inconsistent pattern of results for the five group types across the two age levels and in the different activities that made up the tasks. Although this inconsistency means that we do not end up with neat answers for the questions posed earlier, it does indicate that we need to rethink our assumptions, which have generally been based on research carried out in other contexts. It also means that we need to question assumptions based on generalisations formed by very limited sets of data.

IMPLICATIONS FOR PRACTICE

Overall, my study shows that the gender composition of a group is not a salient factor in children's task groups when they have the opportunity to focus on a shared goal requiring input from the different group members. As Maccoby (1998) suggests, goal-focused work is an important means of bringing boys and girls together, for in such situations the goal becomes an important uniting factor that overrides and reduces the salience of gender issues. This study suggests that individual differences between children are more important than gender in determining their contribution to and participation in group work.

My study also shows that the discourse which continues to suggest that females 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. As Maccoby (1998, *viii*) aptly observes, 'the social context of gender issues ... has been changing and any writing on this subject must be thought of as a work in progress'. As the social context changes, teachers in particular and educators in general need to ensure that they do not perpetuate assumptions and assertions which no longer hold true. This study challenges the almost stereotypical belief



that boys and girls cannot work together effectively.

Moreover, this study also shows that children's experience in same-gender groups is not necessarily more equitable than in the mixed-gender groups, or that it is 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.

Finally, this study shows that having equal numbers of boys and girls in balanced mixed-gender groups does not solve problems inherent in group work. The group experience usually was no more equitable in the balanced mixed-gender groups than in the imbalanced mixed-gender groups. Teachers need to be aware that equal numbers of boys and girls in each group does not necessarily result in an equitable experience, and that this group structure does not protect children (boys as well as girls) from becoming sidelined in the activity.

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. After all, being the minority in a group setting is one of life's realities with which children need to learn to cope.

Within my study, the process of cross-coding the patterns of behaviour and achievement exhibited by children in a relatively large number of groups working on different tasks at two age levels demonstrated that it is impossible to focus on numerous processes accurately during one viewing. Teachers therefore are advised to focus on a limited number of processes at any one time.

Although the performance assessment setting has limitations in terms of being able to extrapolate the results to classroom settings, I believe that the results from this study do have validity for such settings, and more so than do results from other studies conducted in other settings. The conclusions drawn in this study are certainly based on more substantial and diverse data. As such, this project has advanced our knowledge about and understanding of the complex experiences of children's task-focused group work.

REFERENCES

Aries, E.J. (1982) Verbal and nonverbal behavior in single-sex and mixed-sex groups: are traditional sex roles changing? *Psychological Reports*, 51, 127-34.

Bailey, S.M. (1993) The current status of gender equity research in American schools. *Educational Psychologist*, 28, 321-39.

Bandura, A. (1986) *Social Foundations of Thought and Action: a social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Barbieri, M.S. and Light, P.H. (1992) Interaction, gender and performance on a computer-based problem solving task. *Learning and Instruction*, 2, 199-213.

Barnes, D. and Todd, F. (1977) *Communication and Learning in Small Groups*. London: Routledge and Kegan Paul.

Bennett, N. (1994) Co-operative learning. In P. Kutnick and C. Rogers (Eds.), *Groups in School*. London: Cassell, 50-65.

Bennett, N. and Cass, A. (1988) The effects of group composition on group interactive processes and pupil understanding. *British Educational Research Journal*, 15(1), 19-32.

Bennett, N. and Dunne, E. (1992a) Assessing group work. In *Managing Classroom Groups*. Herts: Simon and Schuster.

Bennett, N. and Dunne, E. (1992b) Co-operative grouping. In *Managing Classroom Groups*. Herts: Simon and Schuster.

Berger, J., Cohen, B.P. and Zelditch, M. (1972) Status characteristics and social interaction. *American Sociological Review*, 37(3), 241-55.

Berger, J., Rosenholt, S.J. and Zelditch, M. (1980) Status organizing processes. *Annual Review of Sociology*, 6, 479-508.

Bossert, S. (1989) Co-operative activities in the classroom. *Review of Research in Education*, 15, 225-50.

Bossert, S.T., Barnett, B.G. and Filby, N.N. (1984) Grouping and instructional organisation. In P. L. Peterson, L. C. Wilkinson and M. Hallinan (Eds.), *The Social Context of Instruction*. Orlando, FL: Academic Press.

Brannen, J. and Brien, M.O. (1995) Childhood and the sociological gaze: paradigms and paradoxes. *Sociology*, 29(4), 729-37.

Burns, R.B. (1994) *Introduction to Research Methods*. Melbourne, VC: Longman Cheshire.

Bush, T.A. (1997) The effect of group composition on achievement and time on task for students completing ILS activities in co-operative pairs. *Journal of Research on Computing in Education*, 30(1), 2-17.

Butler, D. and Geis, F.L. (1990) Nonverbal affect responses to male and female leaders: implications for leadership evaluations. *Journal of Personality and Social Psychology*, 58, 48-59.

Carli, L.L. (1982) *Are Women More Social and Men More Task Oriented?* Amherst, MS: University of Massachusetts.

Carli, L.L. (1989) Gender differences in interaction style and influence. *Journal of Personality and Social Psychology*, 56(4), 565-76.

Claris, C. (1995) *FileMaker Pro* (Version 3.0). California: Claris.

Cohen, E. (1984) Talking and working together: status, interaction and learning. In P.L. Peterson, L.C. Wilkinson and M. Hallinan (Eds.), *The Social Context of Instruction*. Orlando, FL: Academic Press, 171-87.

Cohen, E. (1986) *Designing Group Work: strategies for the heterogeneous classroom*. New York: Teachers College, Columbia University.

Cohen, E. (1994) Restructuring the classroom: conditions for productive group work. *Review of Educational Research*, 64(1), 1-35.

Cowie, H. and Rudduck, J. (1989) What is co-operative group work? In R. McCormick, R. Murphy and M. Harrison (Eds.), *Teaching and Learning Technology*. Wokingham: Addison-Wesley.

Crooks, T. and Flockton, L. (1994) *National Education Monitoring Project—Phase II. Preparation for the Implementation in 1995 of National Monitoring of Educational Outcomes in New Zealand Primary Schools (3)*. Dunedin: Educational Assessment Research Unit, University of Otago.

Crooks, T. and Flockton, L. (1996) *Science Assessment Results 1995* (National Education Monitoring Report 1). Dunedin: Educational Assessment Research Unit, University of Otago.

Crooks, T. and Flockton, L. (1996) *Technology Assessment Results 1995* (National Education Monitoring Report 5). Dunedin: Educational Assessment Research Unit, University of Otago.

Damon, W. (1984) Peer education: the untapped potential. *Journal of Applied Developmental Psychology*, 5, 331-43.

Damon, W. and Phelps, E. (1989) Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13, 9-19.

Davies, B. (1989) The discursive production of the male/female dualism in school settings. *Oxford Review of Education*, 229-41.

Denzin, N. (1970) *The Research Act: a theoretical introduction to sociological methods.* New York: McGraw-Hill.

Dion, K.L. (1985) Sex, gender, and groups: selected issues. In V.E. O'Leary, R.K. Unger and B. S. Wallston (Eds.), *Women, Gender and Social Psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Eagly, A.H. (1987) *Social Differences in Social Behaviour: a social role interpretation*. Hillsdale, NJ: Erlbaum.

Edwards, A. (1994) The curricular application of classroom groups. In P. Kutnick and C. Rogers (Eds.), *Groups in Schools*. London: Cassell.

Envisionology (1993) *CVideo* (Version 1.5). San Francisco: Knowledge Revolution.

Fassnacht, G. (1979) *Theory and Practice of Observing Behaviour.* London: Academic Press.

Fontana, A. and Frey, J.H. (1994) Interviewing: the art of science. In N.K. Denzin and Y.S. Lincoln (Eds.), *Handbook of Qualitative Research*. Great Britain: The Falmer Press.

Galton, M. and Williamson, J. (1992) *Group Work in the Primary School*. London: Routledge.

Good, T.L. and Marshall, S. (1984) Do students learn more in heterogeneous or homogeneous groups? In P.L. Peterson, L.C. Wilkinson and M. Hallinan (Eds.), *The Social Context of Instruction*. Orlando, FL: Academic Press.

Hall, E. (1994) The social relational approach. In P. Kutnick and C. Rogers (Eds.), *Groups in School*. London: Cassell.

Hare, P.A. (1992) *Groups, Teams, and Social Interaction: theories and applications*. New York: Praeger.

Hertz-Lazarowitz, R. (1992) Understanding interactive behaviors: looking at six mirrors of the classroom. In R. Hertz-Lazarowitz and N. Miller (Eds.), *Interaction in Co-operative Groups: the theoretical anatomy of group learning*. USA: Cambridge University Press.

Hitchcock, G. and Hughes, D. (1993) *Research and the Teacher: a qualitative introduction to school based research*. London: Routledge.

Hoffman, R.L. (1965) Group problem solving. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2). New York: Academic Press.

Holden, C. (1993) Giving girls a chance: patterns of talk in co-operative group work. *Gender and Education*, 5(2), 179-811.

Howes, C. and Phillipsen, L. (1992) Gender and friendship: relationships within peer groups of young children. *Social Development*, 1(3), 230-42.

Jacklin, C.N. and Maccoby, E.E. (1978) Social behaviour at thirty-three months in same-sex and mixed-sex dyads. *Child Development*, 49, 557-69.

Johnson, D.W. and Johnson, R.T. (1979) Conflict in the classroom: controversy and learning. *Review of Educational Research*, 49(1), 51-70.

Johnson, D.W. and Johnson, R.T. (1992) Positive interdependence: key to effective co-operation. In R. Hertz-Lazarowitz and N. Miller (Eds.), *Interaction in Co-operative Groups: the theoretical anatomy of group learning*. USA: Cambridge University Press.

Johnson, R.A.and Schulman, G.I. (1989) Gender-role composition and role entrapment in decision-making groups. *Gender and Society*, 3(3), 355-72.

Johnson, R.T. and Johnson, D.W. (1994) An overview of co-operative learning. In J.S. Thousand, R.A. Villa and A.I. Nevin (Eds.), *Creativity and Collaborative Learning*. Baltimore, MD: Paul H Brookes Publishing Co., 31-44.

Johnson, R.T., Johnson, D.W. and Stanne, M.B. (1985) Effects of co-operative, competitive and individualistic goal structures on computer-assisted instruction. *Journal of Educational Psychology*, 77, 668-78.

Kanter, R.M. (1977) Some effects of proportions on group life: skewed sex ratios and responses to token women. *American Journal of Sociology*, 82, 965-90.

King, A. (1992) Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist*, 27, 111-26.

Kutnick, P. (1988) *Relationships in the Primary School Classroom*. London: Paul Chapman.

Kutnick, P. (1994) Use and effectiveness of groups in classrooms: towards a pedagogy. In P. Kutnick and C. Rogers (Eds.), *Groups in Schools*. London: Cassell.

Kutnick, P. (1997) Computer-based problem-solving: the effects of group composition and social skills on a cognitive, joint action task. *Educational Research*, 39(2), 135-47.

Leaper, C. (1991) Influence and involvement in children's discourse: age, gender and partner effects. *Child Development*, 62, 797-811.

Lee, M. (1993) Gender, group composition, and peer interaction in computer-based cooperative learning. *Journal of Educational Computing Research*, 9(4), 549-77.

Levine, J.M. and Moreland, R.L. (1990) Progress in small group research. *Annual Review of Psychology*, 41, 585-634.

Light, P. and Littleton, K. (1994) Cognitive approaches to group work. In P. Kutnick and C. Rogers (Eds.), *Groups in Schools*. London: Cassell, 87-103.

Lindow, J.A., Wilkinson, L.C. and Peters, P.L. (1985) Antecedents and consequences of verbal disagreements during small-group learning. *Journal of Educational Psychology*, 79, 658-67.

Linn, R. (1993) Educational assessment: expanded expectations and challenges. *Educational Evaluation and Policy Analysis*, 15(1), 1-16.

Maccoby, E.E. and Jacklin, C.N. (1987). Gender segregation in childhood. In H.W. Reese (Ed.), *Advances in Child Development and Behaviour* (Vol. 20). Orlando, FL: Academic Press, 239-87.

Mael, F.A, (1998) Single-sex and coeducational schooling: relationships to socioemotional and academic development. *Review of Educational Research*, 68(2) 101-29.

McCloskey, L.A. and Coleman, L. (1992) Difference without dominance: children's talk in mixed- and same-sex dyads. *Sex Roles*, 27(5/6), 241-57.

Meeker, B.F. and Weitzel-O'Neill, P.A. (1977) Sex roles and interpersonal behaviour in task-oriented groups. *American Sociological Review*, 42, 91-105.

Miller, P.M., Dahaner, D.L. and Forbes, D. (1986) Sex-related strategies for coping with interpersonal conflict in children aged five and seven. *Developmental Psychology*, 22(4), 543-48.

Ministry of Education (1993) *The New Zealand Curriculum Framework*. Wellington.

Moss, P. and Petrie, P. (1997) *Children's Services: a time for a new approach: a discussion paper*. London: Thomas Coram Research Unit, Institute of Education, University of London.

Murray, M. (1988) Negotiation and collaboration in problem solving groups. *Education 3-13*, 16, 48-52.

Nastasi, B.K. and Clements, D.H (1991) Research on co-operative learning: implications for practice. *School Psychology*, 20(1), 110-31.

Oakes, J. and Lipton, M. (1990) *Making the Best of Schools: a handbook for parents, teachers and policymakers*. New Haven, MS: Yale University Press.

Piaget, J. (1967) *The Psychology of Intelligence*. London: Routledge and Kegan Paul.

Pitcher, E.G. and Schultz, L.H. (1983) *Boys and Girls at Play: the development of sex roles.* Sussex: The Harvester Press.

Prout, A and James, A. (1990) Re-presenting childhood: time and transition in the study of childhood. In A. James and A. Prout (Eds.), *Constructing and Reconstructing Childhood: contemporary issues in the sociological study of childhood*. London: The Falmer Press.

Pryor, J. (1995) Gender issues in group work: a case study involving work with computers. *British Educational Research Journal*, 21(3), 277-88.

Qualitative Solutions and Research Pty (1995) *QSR NUD*IST (Version 3)*. London: Sage.

Quicke, J. and Winter, C. (1995) 'Best friends': a case study of girls' reactions to an intervention designed to foster collaborative group work. *Gender and Education*, 7(3), 259-81.

Qvortrup, J. (1990) A voice for children in statistical and social accounting: a plea for children's right to be heard. In A. James and A. Prout (Eds.), *Constructing and reconstructing childhood: contemporary issues in the sociological study of childhood*. London: The Falmer Press.

Rogoff, B. (1990) *Apprenticeship in Thinking*. Oxford: Oxford University Press.

Rosenshine, B. and Furst, N. (1973) The use of direct observation to study teaching. In R. Travers (Ed.), *Second Handbook of Research on Teaching*. Chicago: Rand McNally and Co.

Salomon, G. and Globerson, T. (1989) When teams do not function the way they ought to. *International Journal of Educational Research*, 13, 89-99.

Savicki, V., Kelley, M. and Lingenfelter, D. (1996a) Gender and group composition in small task groups using computer-mediated communication. *Computers in Human Behavior*, 12(2), 209-24.

Savicki, V., Kelley, M. and Lingenfelter, D. (1996b) Gender, group composition and task type in small task groups using computer-mediated communication. *Computers in Human Behavior*, 12(4), 549-65.

Schmuck, R.A. and Schmuck, P.A. (1983) *Group Processes in the Classroom*. Iowa: Wcb.

Schunk, D.H. (1987) Peer models and children's behavioural change. *Review of Educational Research*, 57, 149-74.

Shantz, C.U. (1987) Conflicts between children. *Child Development*, 58, 283-305.

Sharan, S. (1980) Co-operative learning in small groups: recent methods and effects on achievement, attitudes and ethnic relations. *Review of Educational Research*, 50, 241-71.

Sharan, S. (1990) Co-operative learning: a perspective on research and practice. In S. Sharan (Ed.), *Co-operative Learning*, New York: Praeger.

Slavin, R. (1980) Co-operative learning. *Review of Educational Research*, 50(2), 315–42.

Slavin, R. (1983) *Co-operative Learning*. New York: Longman.

Slavin, R. (1987) Development and motivational perspectives on co-operative learning: a reconciliation. *Child Development*, 58, 1161–167.

Slavin, R.E. (1990a) *Co-operative Learning: theory, research and practice*. Boston: Allyn and Bacon.

Slavin, R.E. (1990b) Co-operative learning. In C.R. Kutnick and. P. Kutnick (Eds.), *The Social Psychology of the Primary School*. London: Routledge.

Slavin, R.E. (1992) Co-operative learning. In M. Alkin (Ed.), *Encyclopaedia of Educational Research*. New York: American Educational Research Association.

Smith, A.B. (1995) *Incorporating Children's Perspectives into Research: challenge and opportunity*. Paper presented at the New Zealand Association for Research in Education Annual Conference, Palmerston North, 7–10 December.

Smith, A.B. (1998) *Understanding Children's Development* (4th ed.). Auckland: Auckland University Press.

Stallings, J.A. and Mohlman, G.G. (1988) Classroom observation techniques. In J.P. Keeves (Ed.), *Educational Research, Methodology and Measurement: an international handbook*. Oxford: Permagon Press.

Strodtbeck, F. and Mann, R. (1955) Sex role differentiation in jury deliberations. *Society*, 19, 3–11.

Tann, S. (1981) Grouping and group work. In B. Simon and J. Willcocks (Eds.), *Research and Practice in the Primary Classroom*. London: Routledge and Kegan Paul, 43–54.

Thomas, G.T. (1994) 'Discussion in Junior Mathematics: helping one another learn?' Unpublished doctoral thesis, University of Otago.

Thousand, J.S., Villa, R.A. and Nevin, A.I. (Eds.) (1994) *Creativity and Collaborative Learning*. Baltimore, MD: Paul H Brookes Publishing Co.

Tolmie, A. and Howe, C. (1993) Gender and dialogue in secondary school physics. *Gender and Education*, 5(2), 191–209.

Tudge, J. and Rogoff, B. (1989) Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M.H. Bortein and J.S. Bruner (Eds.), *Interaction in Human Development* (Vol. 1). New Jersey: Lawrence Erlbaum Ass. Inc.

Underwood, G., Jindal, N. and Underwood, J. (1994) Gender differences and effects of co-operation in a computer-based language task. *Educational Research*, 36(1), 63–74.

Underwood, G., McCaffney, M. and Underwood, J. (1990) Gender differences in a co-operative computer-based language task. *Educational Research*, 32(1), 44–49.

Vygotsky, L.S. (1962) *Thought and Language*. Cambridge, MS: MIT Press and Wiley.

Vygotsky, L.S. (1978) *Mind in Society: the development of higher psychological processes*. Cambridge, MS: Harvard University Press.

Walkerdine, V. (1981) Sex, power and pedagogy. *Screen Education*, 38(Spring), 14-24.

Wasserman, G.A. and Stern, D.N. (1978) An early manifestation of differential behaviour toward children of the same and opposite sex. *The Journal of Genetic Psychology*, 133, 129-37.

Webb, N.M. (1982a) Group composition, group interaction and achievement in co-operative small groups. *Journal of Educational Psychology*, 74(4), 475-84.

Webb, N.M. (1982b) Peer interaction and learning in co-operative small groups. *Journal of Educational Psychology*, 74(5), 642-55.

Webb, N.M. (1982c) Student interaction and learning in small groups. *Review of Educational Research*, 52(3), 421-45.

Webb, N.M. (1983) Predicting learning from student interaction: defining the interaction variables. *Educational Psychologist*, 18(1), 33-41.

Webb, N.M. (1984) Sex differences in interaction and achievement in co-operative small groups. *Journal of Educational Psychology*, 76(1), 33-44.

Webb, N.M. (1989) Peer interaction and learning in small groups. *International Journal of Educational Research*, 13, 21-40.

Webb, N.M. (1991) Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, 22(5), 366-89.

Webb, N.M. (1993) Collaborative group versus individual assessment in mathematics: processes and outcomes. *Educational Assessment*, 1(2), 131-52.

Webb, N.M. (1994) *Group Collaboration in Assessment: competing objectives, processes and outcomes* (CSE Technical Report 386). Los Angeles: CRESST/University of California.

Webb, N.M. (1995) Group collaboration in assessment: multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis*, 17(2), 239-61.

Webb, N.M. (1995) Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology*, 87(3), 406-23.

Webb, N.M. (1997) Assessing students in small collaborative groups. *Theory into Practice*, 36(4), 205-13.

Webb, N.M. and Kenderski, C.M. (1984) Student interaction and learning in small-group and whole-group settings. In P.L. Peterson, L.C. Wilkinson and M. Hallinan (Eds.), *The Social Context of Instruction*. Orlando, FL: Academic Press.

Webb, N.M. and Kenderski, C.M. (1985) Gender differences in small group interaction and achievement in high- and low-achieving classes. In L.C. Wilkinson and C.B. Marrett (Eds.), *Gender Influences in Classroom Interaction*. Orlando, FL: Academic Press.

Webb, N.M. and Palinscar, A.S. (1996) Group processes in the classroom. In D. Berliner and R. Calfee (Eds.), *Handbook of Educational Psychology* (3rd ed.). New York: Macmillan.

Weiler, K. (1988) *Women Teaching for Change: gender, class and power.* Massachusetts: Bergin & Garvey Publishers, Inc.

Whiting, B.B. and Edwards, C.P. (1988) *Children of Different Worlds: the formation of social behaviour.* Cambridge, MS: Harvard University Press.

Wilkinson, L.C., Lindow, J. and Chiang, C.P. (1985) Sex difference and sex segregation in students' small-group communication. In L.C. Wilkinson and C.B. Marrett (Eds.), *Gender Influences in Classroom Interaction.* Orlando, FL: Academic Press.

Willems, J. (1981) Problem-based (group) teaching — a cognitive science approach to using available knowledge. *Instructional Science*, 10(1), 5-21.

Wood, W. (1987) Meta-analytic review of sex differences in group performance. *Psychological Bulletin*, 102(1), 53-71.

Yeomans, A. (1983) Collaborative group work in primary and secondary schools: British and the USA. *Durham and Newcastle Research Review,* 10(51), 99-105.

APPENDIX

MARKING SCHEDULES

SEPARATING MIXTURES

PLANNING — WITHOUT KNOWLEDGE OF EQUIPMENT

For questions 1 to 3 use this rating scale:
4: Good 3: Moderate 2: Weak 1: Absent

1. Evidence of systematic approach.
2. Consideration of useful equipment.
3. Qualities of the material discussed.

PLANNING — WITH KNOWLEDGE OF EQUIPMENT

4. Discuss uses of equipment in relation to materials: Yes No
5. Modification of plan to incorporate use of equipment: Yes No

For questions 6 & 7 use this rating scale:
5: Excellent 4: Very Good 3: Good 2: Fair 1: Weak

6. Overall planning score.
7. Team work in relation to planning.

EXPERIMENTATION

For questions 8 – 10 use this rating scale:
3: Good 2: Moderate 1: Weak

8. Evidence of systematic approach.
9. Effective use of equipment.
10. Team work.

REPORT TO TEACHER

11. Student's evaluation of their plan:

3: Very good: clear and accurate as to strengths and weaknesses.
2: Moderate: have some appreciation of strengths andweaknesses.
1: Weak: little idea of relevance of plan.

QUESTION TIME

1. Have provided a range of questions which are relevant to topic: Yes No
2. Have identified which six they wish to present to the police officers: Yes No

For questions 3.1 – 3.6 use this rating scale:
1: Questions requiring a one or two word answer.
2: Questions requiring a sentence or so.
3: Questions giving opportunity for an extended answer.

3.1 First question for police.
3.2 Second question for police.
3.3 Third question for police.
3.4 Fourth question for police.
3.5 Fifth question for police.
3.6 Sixth question for police.

SPACE GAME

1. Quality of ideas selected for making the game more fun:
1: Weak 2: Moderate 3: Strong

2. Quality of plan to find out if other children think the game is fun and to collect ideas for improving the game:

1: Low
2: Moderate
3: Quite high
4: Very high