WORKING TOGETHER IN THE CLASSROOM: AN INVESTIGATION INTO SOFTWARE TO RAISE AWARENESS OF GROUP-LEARNING SKILLS IN CHILDREN AGED 9 AND 10

by

MARY HELEN ULICSAK

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Electronic, Electrical and Computer Engineering
The University of Birmingham
August 2002

Pages 0 to 219
Main body of thesis

Appendices are in an additional file

ABSTRACT

This thesis describes a training scheme to raise awareness of group skills. Drawing on existing research and observations of classroom practice, this scheme requires children to consider group skills before, during and after the group activity. The approach is unique as the task of raising awareness of group skills was designed to influence, *and* be influenced by, the group task; and feedback on group skill usage is generated from individual self-assessments made during and after the activity.

Studies using the scheme with 9 and 10 year olds working in face-to-face groups found the medium (paper or software) and the environment (a classroom or laboratory-style setting) influenced the self-assessments. Despite this, using the scheme did not result in the transfer of group skills to other activities. A relationship was found, however, between the consistency of the self-assessments recorded during and after the activity and the child's ability at group work.

The training scheme positively influenced the task performance. If a software implementation of the scheme was used concurrently with computer tasks, there was a trend towards improved recall in the material studied. If used with an activity that also incorporated communication, reflection and responsibility, a significant improvement in performance was achieved.

Keywords

Group work, computer supported collaborative learning

ACKNOWLEDGEMENTS

I wish to thank the pupils and teachers from the schools who allowed me to observe then disrupt their lessons. In the first study, this was Bleakhouse Junior School, St Joseph's Primary School and The Shrubbery. For the remaining studies, I am grateful to St Peter's Junior School and Alston Junior School. These two schools allowed me to not only take the children out of classes but rescheduled timetables and groups so that I could run the Information and Communication Technology classes. Without the help and support of these teachers none of this work would have been possible. I am also indebted to the technicians in the University of Birmingham and in the schools for ensuring I had all the recording devices and computers necessary.

I would like to acknowledge the support given by my supervisors, Professor Harry Daniels and Professor Mike Sharples. In addition, I am grateful to the following for their advice, support and proof reading: Dr T Arvanitis, Andy Bracher, Trevor Francis, Dom Maxted, Liz Masterman, Sarah Whitehead and Tom Ulicsak. Also to Victoria Bell, for the extended loan of her laptop computer.

Most of all, however, I wish to thank Alex Gibb, whose unreserved generosity in time, infinite patience and common sense I have relied on throughout my time in Birmingham.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF BOXES

LIST OF TABLES

CONVENTIONS

The pupil is referred to as "he".
The teacher is referred to as "she".

In the transcript extracts:

{}	The group members are talking simultaneously
|	The next speaker interrupted the sentence
…	Pause in conversation

To ease comprehension for the reader the transcripts have been punctuated.

The studies in this research followed The British Psychological Society guidelines (2000).

Chapter 1

INTRODUCTION

This thesis describes the development and testing of a teaching scheme to raise awareness of group skills within children aged 9 and 10 working in a classroom setting. In particular, it investigates the impact of using computers as a medium for the scheme. This chapter summarises the approach taken and the methodology used, and concludes with an overview of the structure of the dissertation.

1.1 Executive summary

This section summarises the goals of the research project, the assumptions made, and the main findings.

1.1.1 The research project

Group work has been shown to benefit children's cognitive and social skills (e.g., Phelps and Damon 1989 , Slavin 1995, Webb 1997, Johnson *et al.* 1990), and is compulsory in the state education system (DfES 2000). Children in Key Stage 2 – that is those aged 7 to 11 – are required to take turns, ask relevant questions, respond appropriately, plan, predict and evaluate within a group discussion. However, the skills required for group work – such as listening, negotiating, planning, justifying, and questioning – are not instinctive (Crook 1998, Biott and Easen 1994d), and so teaching schemes exist to develop them.

Such schemes provide a structure for the interactions among group members. They require the children to display three interrelated attributes: communication, responsibility and reflection. The term communication refers to the set of skills needed for information and individual understandings to be expressed. It is a combination of exploratory talk, the offering of timely explanations, listening to others, asking questions, acknowledging the contribution of others (Mercer 1994) and non-semiotic exchanges, such as gestures, pauses and eye contact. The second attribute, responsibility, must be towards the individual *and* the group, that is the child is responsible for ensuring that both he and the others in the group understand the task. Reflection allows the development of ideas, predictions and evaluations. It is the process of internalisation and knowledge construction and the exchanges could be an external or internal dialogue (Säljö 1998). All three attributes require language. Indeed, Vygotsky (1978) states that speech is the

method through which we organise our thinking, regulate our actions and acquire the culture of the society in which we live.

The training scheme developed in this project proposed that a group activity be viewed as two simultaneous tasks. The first is the teacher-given task (TGT), which can be directed either towards a specific curriculum topic (for example, conducting an experiment in order to understand factors affecting plant growth), or towards a specific object (for example, producing a class newspaper). The second is the group skills task (GST). The GST raises awareness rather than directly teaching the skills that allow the group members to work effectively together. Drawing on existing training schemes and observations of classroom practice, the GST requires the children to consider group skills *before*, *during* and *after* the group activity. The Developing and Supporting Group Skills (DSGS) model was devised to scaffold the children through these stages. First it requires the children to establish: (i) rules for working together for the TGT, and (ii) criteria to assess them (Farivar and Webb 1994, Ashman and Gillies 1997, Mercer *et al.* 1999). During the TGT, the children are prompted to *individually* assess the amount of time they believed that they used one of these specified group skills using the agreed criteria. On the basis of these opinions, the DSGS model then generates discussion questions (Palincsar *et al.* 1993, King *et al.* 1998) and randomly assigns roles associated with specific cognitive skills and types of behaviour to the group members (e.g., Goodman *et al.* 1998). After the task, the children individually assess how well they worked together overall. Finally, using both types of self-monitoring data the model generates feedback in the form of a numeric value corresponding to the time spent using group skills and an appropriate discussion prompt. The numeric value correlated with the opinion of teachers observing group skill usage in the pilot sessions.

The approach is novel for two reasons:

- *There is an interaction between the TGT and the GST.* The children's behaviour during the TGT determines the type of discussion questions generated by the GST, and the answers to those questions should influence subsequent behaviour in the TGT
- *Feedback, which is necessary to ensure that the children's understanding of group skills and assessment of usage within the GST is reasonable, is based almost solely on the children's own assessment of their behaviour.* In previous research it required an analysis of the level of interactions, amount of help requested, discourse or final product (Ager 1998, Dunne and Bennett 1996, Wild and Braid 1996, Gillies and Ashman 1998, Mercer *et al.* 1999)

It was hypothesised that by making the children explicitly aware of group skills and asking them to consider and discuss their behaviour as part of the GST this would positively influence the interactions within the TGT. That is, the TGT would result in an improved product and greater individual understanding. The prompting process by the GST would act as a simplistic form of scaffolding (Wood *et al.* 1976), since it models the type of questions, turn taking and justification that should be used in the TGT. In turn, the feedback from the GST would be more positive as the children had used group skills throughout the TGT. This hypothesis was tested by determining if:

- The children performed the group task more effectively; judged by a better product, a more elegant solution, or whether individual children recalled more information after studying as a group
- These group skills were transferred to other TGTs where there was no simultaneous GST, that is whether children had formalised their everyday skills into scientific or scholarly ones (Vygotsky 1987, Kozulin 1986)

1.1.2 Assumptions

On the basis of existing research, four assumptions were made during the design phase and tested as part of the research programme. These are:

- A training scheme that explicitly supports communication, responsibility and reflection will raise awareness of group skills (Johnson *et al.* 1990)
- Context is influential. In terms of activity settings, the training scheme acts as the mediating artefact. The classroom environment, perceptions of appropriate behaviour, existing group skills, the children, the teacher, the expectations of feedback and the history of each of these will impact its effectiveness (Engeström 1999a)
- Children in the target age group have sufficient metacognitive ability to perform the TGT and the GST in parallel. The GST requires the children to reflect on their own behaviour, although the training scheme scaffolds this process by asking them to focus on one skill over a few minutes using defined criteria (Azmitia 1998¸ Brown 1987)
- The computer is an appropriate medium for developing group skills as it encourages discussion (e.g., Light 1993, Hoyles *et al.* 1991b, Wegerif 1996b)

1.1.3 Outcome of the research

The training scheme appeared to influence the amount of information retained when 9 and 10 year olds in Year 5 worked in pairs at a computer. In the second study conducted the children were given a TGT that explicitly required communication, responsibility and reflection. Performing this task *and* being simultaneously exposed to the DSGS model resulted in a significantly better performance by the children than if just the TGT was given. In a later study, where the TGT did not explicitly incorporate these three attributes, those pupils who were simultaneously exposed to the model demonstrated a better individual recall of information learned within the TGT than those who were not.

The feedback from the computer was intended to provide a starting-point for pupils to reflect on their behaviour and in so doing *appropriate* (Newman 1990) the concepts associated with group skills so that they would transfer them to other activities. In practice though, the children treated the feedback as the end-point of the task and did not engage in the reflective activities suggested. However, despite not raising awareness of group skills within the children, the software appears to have potential for identifying those children who are developing group skills. A correlation emerged between the self-assessment data recorded during the TGT and the data captured after it for around 60% of the children. In independent assessments conducted without knowledge of this data around 76% of those who had improved in the teacher's judgement were more consistent. Furthermore, of those children felt not to improve in group skill usage over 80% were more inconsistent in self-assessments.

It was also found that the medium (computer or pencil and paper) and context (classroom or laboratory-style setting) influenced the self-assessments. Compared to paper when using a computer as the medium for the DSGS model the children were more likely to give the same self-assessments, believe they used a group skill a greater proportion of the time during the TGT and give a lower overall mark for group skill usage. They spent more time on the GST and were more consistent in the self-assessments during the task and at the end of the TGT under observation in a more formal laboratory-style setting. It appears, then, that in children aged 9 and 10 the metacognitive ability needed to make realistic self-assessments is affected by the use of a computer or if they are in a classroom environment.

One possible explanation is that pupils who are used to working with "drill-and-skill" mathematics packages on the computer come to believe that the computer expects the "right" answer. Thus in the studies described they were more likely to confer with each other in order to achieve this outcome. In the classroom environment, the children may believe that the TGT takes priority, while in the laboratory-style setting, they were aware of the researcher who could prompt at each stage. Even if no direct input was given the discussions highlighted misconceptions about groups and group skills held by the children, and led to them constructing shared definitions. This suggests that to be effective a training scheme based on the DSGS model requires input from a teacher at all stages.

1.2 Methodology

This section describes the approaches taken and the limitations of the research project.

1.2.1 Format of the research project

The approach taken in this research project followed the principles of Action Research (Cohen and Manion 1998, p.186), which comprises two stages: diagnostic and therapeutic.

The diagnostic stage began with a literature review to establish what training schemes were advocated in order to raise an awareness of group skills. Situated studies were then conducted in local primary schools to:

- Confirm that the skills of responsibility, reflection, and communication had been correctly identified
- Make recommendations for a training scheme to raise awareness of group skills that could be integrated into existing classroom practice
- Investigate whether the computer is a suitable medium for the training scheme, and if so, suggest design criteria

The studies were necessary since observations and analysis of classroom activities are prerequisites to designing a system that promotes learning (Edwards and Mercer 1987, p.17). The concept of subdividing a group activity into the GST and TGT emerged from the diagnostic stage and became the basis of the DSGS model.

In the therapeutic stage the DSGS model was tested both inside and outside the classroom environment. An iterative set of studies was performed in order to refine the model (Boehm 1988). Additional studies examined the impact of the medium in which the model was implemented and the feedback given to pupils.

All the software was designed with the aid of child "informants" (Druin 2002). They served as users and testers. They also assisted in the construction of tools to measure the impact of using the training scheme. Both quantitative, e.g., task performance, and qualitative data, e.g., discourse analysis, were required as neither is sufficient to represent the understanding of an individual within a group (Fisher 1993, Howe 1991). Teachers also participated in the design process, reviewing the wording, format, and in the case of the software used within the intervention program, that the domain content matched National Curriculum requirements.

The studies were designed to incorporate triangulation (Cohen and Manion 1998, pp.233-251). Within each study, quantitative and qualitative data were gathered and used to support each other. Examples include: contrasting the children's performance with their spoken opinions in the situated studies (Study 1a) and comparing the recorded self-assessments using the software with the questionnaire opinions (Study 2). Between-study corroboration included testing the impact of the setting (classroom or laboratory-style) and the medium in which the DSGS was implemented (paper or computer) on the children's behaviour.

Once the model had been refined, a quasi-experimental approach was taken to test the hypothesis that the DSGS model would improve cognitive and social skills within group activities. It cannot be classed as a true experiment as there was no random selection of schools; the research was restricted to those willing to participate (Cohen and Manion 1998, p.169). The study was designed so that the hypothesis would be tested in an ecologically valid environment, since using a laboratory-style setting would pose the risk that the activity would be seen as a novelty and hence distort the findings (Crook 1991). It was expected that implementing the DSGS model in the classroom as an integral part of teaching practice would highlight whether the computer could overcome the boundaries surrounding it: for example, the attitudes of teachers and pupils, and the impact of the feedback given by the computer. This is necessary as in Activity Theory the action cannot be divorced from the context (Kuutti 1996).

The intervention programme followed a "pre-test, system use, post-test" format, with experimental and control groups taken from two Year 5 classes. The independent variable was a programme of 18 ICT lessons in which the experimental group were exposed to the DSGS model and the control group were not. The dependent variables were: the children's discourse, the teachers' assessment of the children usage of group skills, the problem-solving ability of each group, and individual pupils' recall of the information studied. The experimental and control groups were matched in order to minimise confounding variables such as teaching style. However, despite this careful experimental design, it is not possible to attribute changes in the children's group skill usage to the DSGS model. This is because the experiment constitutes a "complex intervention" (MRC 2000), the outcome is dependent on the interrelation of additional, uncontrollable factors such as the composition of the group, the type of activity, the time of day when the activity is performed and the fact that children can have "off days" etc. (Dunne and Bennett 1996).

1.2.2 Limitations of the methodology

Ideally a true Action Research approach would have been employed. As Galton and Williamson (1994, p.170) argue, effective teaching strategies have to be adopted by the teacher instead of being imposed on them. Training schemes already in use – for example, co-operative integrated reading and composition (CIRC), Jigsaw activities, student-teams achievement division (STAD) and the Open University Talk Lessons – have a high success rate due to the teachers' commitment to collaborative learning (Slavin 1996, Bennett and Dunne 1992a, Dawes *et al.* 2000). In contrast, the teachers involved in these studies acknowledged the benefits of group work but did not adapt their teaching style; they did not act as co-researchers and contribute to the development of the DSGS model. They did, however, actively support the research by participating in interviews and designing the software, reviewing videos of children working together, working on a formal set of actions that represented group skill usage, completing questionnaires and assessing how well children worked together in ICT lessons.

As stated in Section 1.2.1 a quasi-experimental approach was used, since the studies were limited to those schools willing to participate. This is a restriction as all five schools participating focused on the transmission of domain knowledge - a collaborative culture was not encouraged in any of the 12 classes involved (Daniels 2001b). Due to this focus, the intended age group for this research project, Year 6, could not participate in Study 2 onwards as the schools wished them

to concentrate on the Standard Assessment Tests (SATs). These older children would be more likely to demonstrate metacognitive skills (Azmitia 1998, Sharples 1999). This was illustrated by the answers given by children from Year 6 in Study 1, who found it easier to take on the perspective of others.

1.3 Overview of thesis

The thesis is divided into four parts. The first examines the theoretical basis of the research; the second part describes the implementation, refinement and testing of the DSGS model; the third part discusses the intervention programme; and the final part summarises the findings, assesses the validity of the assumptions made, and proposes areas for future research.

1.3.1 Diagnostic stage: Theoretical basis of the model

Chapters 2 and 3 cover the diagnostic stage of the project. Chapter 2 begins by elaborating on why group work is used within the classroom, and the pedagogical basis behind the types of groups that could be used. It also looks at the development of the computer as a mediating artefact for groups.

Chapter 3 describes Study 1, the situated study used to establish current classroom practice and perceptions of group work. The goals of the study were to establish that the attributes for effective group work had been correctly identified, suggest requirements for the training scheme, and investigate whether a computer would be a suitable medium.

1.3.2 Therapeutic stage: Development of the model and teaching scheme

The first version of the DSGS model is presented in Chapter 4. This chapter also covers Study 2, which tested the impact of the model on a TGT in the form of a computer puzzle. The findings of this study – a limited level of self-monitoring by pupils, together with an overestimation of overall group skill usage compared to the teachers observing them – led to the revisions of the model which are described in Chapter 5, together with the tests carried out on those revisions (Studies 3a and 3b).

In Chapter 6, the focus switches to the impact of the medium on pupils' assessment of their group skill usage. Study 4a compares the impact of an implementation of the DSGS v2 model

either as a computer system or as a paper-based scheme in a laboratory-style setting. Study 4b is a similar comparison but within a classroom.

Chapter 7 describes the theoretical basis of the heuristic for assessing the amount of group skill usage from the children's self-assessments and the practical issues of a computer providing feedback on the children's ability to work together (Study 5).

1.3.3 Quasi-experimental study: Intervention programme

Chapter 8 describes the impact of 18 weeks' usage of a training scheme, based on the DSGS v2 model (Study 6). This intervention study was required since Study 1 showed a single session does not result in raising the children's awareness of group skills, it is only through repeated use that a child is likely to become socialised and trained in the activity – and as a result innovate and change their understanding.

1.3.4 Review and further work

In the final part of the thesis conclusions are drawn. Chapter 9 reviews the outcome of the research. It then assesses the validity of the four assumptions that were drawn from the design phase and tested as part of the research programme. These concern: the attributes required within a training scheme to raise awareness of group skills, the appropriate audience, issues for evaluation, and the appropriate medium for the training scheme. The contributions of this research are then summarised. Next areas for future work generated from the existing findings and related research are considered. They address the reliability of the current research programme findings, improvements to the DSGS models, and refinements to computer implementations of the training scheme.

Chapter 2

GROUP WORK, CLASSROOMS, AND COMPUTERS

The ability to work in groups is a "life skill" (Johnson *et al.* 1990, p.88). Group skills, such as listening, providing justifications, taking turns, and negotiating, are used regardless of the group's goal (Farivar and Webb 1994, Ashman and Gillies 1997, Wegerif and Dawes 1998, Bennett and Dunne 1992e, etc.). These skills, however, are not instinctive (e.g., Biott and Easen 1994d, Crook 1998). As Johnson *et al.* state: "There is a crucial difference between simply putting students into groups to learn and in structuring cooperation among students" (1990, p.10). The goal of this research program is to develop a training scheme that raises awareness of group skills among children aged 9 and 10. It is hypothesised this will result in:

- The children performing the group task more effectively, that is, they produce a better piece of work, a more elegant solution, or individually recall more information when studying as a group
- Children formalising everyday skills into scientific or scholarly ones by explicitly considering their interactions. This formalisation allows them to apply these skills to other group activities

This chapter discusses the theoretical basis for such claims. It looks at why group work is used, and what type of groups can exist within a primary school classroom. Through an analysis of these types, it proposes there are three attributes necessary for working effectively together: communication, responsibility and reflection. These are elaborated upon, and corroborated by examining the focus of existing training schemes. The training schemes are dependent upon the activity setting. Consequently, the role of groups within the classroom is discussed. The context also includes the medium for the training scheme. In this research program the role of computers as a medium is a focal area. The use of computers within primary schools is therefore examined.

2.1 Why group work is used

It has been found that group working in the classroom promotes individual critical thinking skills, including the ability to analyse, evaluate, synthesise, and apply information (Reid *et al.* 1989, pp.11-12, Forman and McPhail 1993, Bennett 1994, Kutnick 1994). Post-tests after group work often show that the group members have a greater task understanding and produce a better final product than if they had worked alone (e.g., Hoyles *et al.* 1991a, Webb 1997, Wegerif *et al.* 1998). Groups can motivate, empower and provide a forum for practicing negotiation and planning skills (Phelps and Damon 1989, Biott and Easen 1994c). They can encourage an appreciation of different perspectives, and tolerance of others (Biott and Easen 1994a, Slavin 1995). For the individual working in groups can result in them being: less critical of themselves, less withdrawn or self-conscious, less influenced by criticism, and less prone to develop psychological problems such as anxiety or depression (Johnson *et al.* 1990, p.38, Biott and Easen 1994b).

Dillenbourg (1999) suggests groups are beneficial as they require interactions, for example, explanations and mutual regulation that may not occur when working individually. In turn these trigger cognitive mechanisms, such as knowledge elicitation – having to describe what they are doing and why. It is also argued that interaction benefits social skills. Johnson *et al.* (1990, p.22-23) state that, through socialisation with peers, children are given models of behaviour and these influence their attitudes and actions.

The cognitive and social benefits are recognised by the English National Curriculum (DfES 2000). Group work must be used within the primary classroom and children are required to reach Attainment Targets for group skill usage. In all subjects, pupils should "use language precisely and cogently" and "listen to others, and to respond and build on their ideas and views constructively" (DfES 2000). The current Key Stage 2 requirements state that a child must demonstrate being an effective member of a group through taking turns, varying their contributions, justifying their opinions, etc. The lowest Attainment Target for this age requires children to "listen to others and usually respond appropriately", and to "extend their ideas or accounts by providing some detail" (DfES 2000). (Appendix A contains the requirements in full.)

Group work, however, is not always the ideal method of teaching and learning, or "the guided construction of knowledge" (Mercer 1995b). Depending on the task and context, children may learn more efficiently by imitation or through rote learning and repeated drills (Azmitia 1998). If a teacher does decide to use groups, her choice is also influenced by the task and context. As discussed in the next section various types of groups, each with a pedagogic basis, exist within a classroom (Galton and Williamson 1994, Bennett and Dunne 1992a).

2.2 Group classifications

In a classroom there are many examples of group tasks; they can be from informal discussions of an activity through to formal organised group work (DfES 2000). All these activities are termed group work but have different attributes and different intended outcomes. Children, for example, may produce individual or shared products, or they may all have to understand every aspect of the activity or only the part assigned to them (Galton and Williamson 1994, pp.6-10). In establishing a training scheme to raise an awareness of group skills so that the group is more efficient, and the children formalise their understanding, it is necessary to understand the theories behind learning in groups and the attributes they require of the participants.

Depending on the outcome desired, groups have been analysed with respect to their formation. The catalyst for learning can be a group, but the process is internal to the child (Doise 1990). Alternatively, learning has been described as first occurring between the group members using various tools and mediating artefacts before being internalised (Vygotsky 1978). Finally, it has been seen as the development of a jointly constructed solution in which the dialogue is equivalent to a single account (Säljö 1998). Dillenbourg *et al.* classified these views on a theoretical axis. At one end groups can be viewed as "relatively independent cognitive systems which exchange messages" and at the other "a single cognitive system with its own properties" (1996, p.190). The advantage of this approach is that it shows the relationship between the three distinct theoretical approaches about groups. The s*ocio-constructivists* who view individual development in the context of social interaction are on the left of the scale. *Socio-culturalists,* who believe that the individual cognitive change is influenced by the context and vice versa, fall in the middle. Whilst those with a *shared or distributed cognition approach* see the group and the context as an integrated whole. This is illustrated in Figure 2.1.

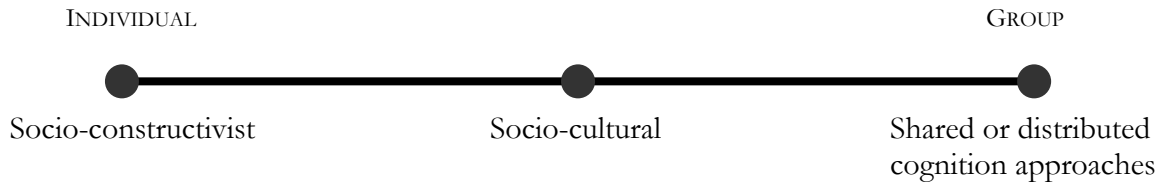


Figure 2.1: Theoretical axis for the unit of analysis of group work based on Dillenbourg et al. (1996, p.190)

As Dillenbourg and his colleagues argue (1996), no theory is better than another; instead they take a different perspective about the unit of analysis ranging from the individual to the group. In the remainder of this section each approach and its use within the classroom is described.

2.2.1 Theoretical approaches to group work

The socio-constructivist approach is based on Piaget's theories of individual cognitive development. Groups are seen as composed of individuals with new understanding being mastered through interactions with others (Doise 1990). In group activities a child is believed to learn through re-evaluating his understanding after finding discrepancies between his thinking and that of his peers. This discovery and resolution of differences has been termed "socio-cognitive conflict". However, the resolution of a problem is also influenced by the context. The children's knowledge of the social norms and rules, for example, meant they could copy a classroom model layout more accurately than a fictional village (Doise 1990). It has also been argued that it is not the disagreement that facilitates change but the communication it generates within the group (Dillenbourg *et al.* 1996). Improved task performances in groups have been ascribed to the resolution of conflict in children aged 7 and 8 (Light and Glachan 1985), 8 to 10 (Barfurth 1995) and 12 to 14 (Howe 1991).

The *socio-cultural* view is in the middle of Dillenbourg *et al.*'s scale (1996). It emphasises that cognitive development is located in sociocultural practices. The participants' attitudes, goals, understandings and actions, are constrained by: the physical and institutional context of an activity, the social roles and the status of the individuals involved, the cultural mediators available, and the prevalent cultural values and beliefs (Daniels 2001a). Vygotsky was the major instigator of this approach (Light and Littleton 1999a, p.xvi). Learning first occurs on the social or *interpsychological* level - it is between the learner and other people using psychological tools such as language and cultural artefacts, such as blackboards and building blocks. This knowledge is then appropriated so that it occurs on the individual or *intrapsychological* level (Vygotsky 1978). This

does not imply that meanings are internalised directly, but that the individual constructs his interpretation, or misconception, from the context in which they exist. Each child understands the actions of others according to his conceptual framework (Newman 1990). Even when working alone socio-culturists argue that learning is still a joint activity, because the child's behaviour is influenced by the instructions he has been given (Mercer 1995b). The process is not one-way, with knowledge being unchanged and passed on directly. Instead, new languages and tools are developed through the interactions (Engeström 1999b). Furthermore, in this type of interaction the children can reconceptualise the given script, that is, the existing rules for interacting (Engeström *et al.* 1997). Here the more able children may play a larger role in influencing and interpreting both of these elements through discussion and reflection (Daniels 2001b).

The s*hared or distributed cognition* approach states that cognitive development occurs within a social structure. The artefacts, tools and group are treated as a single unit, with knowledge being *co-constructed*. The knowledge within the group members can be distributed. The example given by Hutchins (1995) is of a ship, navigation is achieved through each person and artefact performing different activities but with the same goal of ensuring the vessel is on course. The members, however, need not have knowledge of how the other individual tasks are performed. In contrast to the idea of each knowing only their own role, distributed cognition has also been used to describe activities where each member understands all aspects of the task. This requires a process of individual and joint reflection, with knowledge being developed by the elaboration and integration of the perspectives of the different participants into a shared view. Edwards and Mercer (1987, p.3) state that it is possible that through communicating two people will achieve a new understanding. This knowledge is not transferred as in the socio-cultural view, or developed internally as believed by the socio-constructivists. Instead, all group members construct the knowledge, and are responsible for their understanding and that of the other group members. As part of this activity there are no defined roles such as explainer and listener (Baker and Lund 1996, Säljö 1998). Solutions are found through communication: "it is neither conflict nor co-operation that is important in collaborative learning but a combination of the two in a form of interaction which encourages critical challenges within a co-operative search for the best solution" (Wegerif 1996b, p.51).

2.2.2 An analysis of groups in the classroom

The types of group work in a classroom can be divided into: *individual activities–individual outcome*, *individual activities–joint outcome*, *joint activities–joint outcome*. This section relates these categories to the scale used by Dillenbourg *et al.* (1996).

The first, individual activities–individual outcomes can be found in *working groups* or *parallel group work* (Galton and Williamson 1994, pp.6-10, Forman and Cazden 1985). In this type of group activity the children work at the same table and are given the same task but are required to produce individual results. The children have no responsibility towards the understanding of others, but through comparisons or discussions about the activity clarify their own understanding. This style of group work has practical advantages. Such groups allow the streaming of pupils and instruction of a group of the same level. In a classroom, for example, this type of group work would be completing the same worksheet. Socio-constructivist groups can be seen as having individual activities–individual outcomes. This because the individual resolves his own differences and the effectiveness of the group is measured in terms of an individual's performance at the task after the activity.

Working group activities are not limited to individual activities–individual outcomes however. Such an activity may fall in the middle of the theoretical axis in Figure 2.1 if one child takes the role of tutor and explains the activity to the others. Or if no group member understands, the children may confer to construct an understanding which means the activity would be classified as collaboration and be represented as distributed cognition on the right of the scale.

The individual activities–joint outcome type of group working includes *co-ordinated* and *co-operative* tasks (Engeström *et al.* 1997, Galton and Williamson 1994, pp.6-10). These belong to the *working as groups* class, which are identified by a single outcome despite being formed of many subtasks (Daniels 2001b). In both co-ordinated and co-operative groups the students are responsible for the individual subtasks to achieve the group goal. In co-ordinated groups these tasks are assigned. In co-operative tasks, interactions between the group members exist, such as the negotiation of roles. In both cases, the children have individual accountability. They are responsible for the task assigned to them but need not have any understanding of how the other subtasks were achieved. The task participants all accept the tacit script for working together, that is, the given rules of engagement accepted within that environment. An example of an individual

activity–joint outcome would be the creation of a class newspaper where one child wrote the headlines, another did the pictures, a third was responsible for the layout etc. All the roles are required for the product, but the person doing the pictures need not understand how the computer software to layout the material works. This is one instantiation of shared or distributed cognition and would fall on the right of the Dillenbourg *et al.* scale (1996)

The third type of group work found in the classroom is joint activities–joint outcome. As in the previous types, each child takes responsibility for their own understanding, but they are also responsible for understanding of other group members. This type of group work also falls into the category of *working as groups* because there is one product. It is distinguished from individual activities–joint outcome as *all* the individuals must understand and be able to perform *all* aspects of the task. There is a positive interdependence between the group members (Johnson *et al.* 1990, p.19). Within a classroom this style of working occurs in *collective* (*peer tutoring*) groups. In these activities, the flow of information between the group members is unequal - some children take responsibility for tutoring others (Vygotsky 1978, p.86). Returning to the example of a class newspaper, if one child was responsible for teaching the other how to use the computer software to layout the materials so that both could achieve the task at the end, they would be working collectively. This falls in the middle of the Dillenbourg *et al.* (1996) scale.

Another example of a joint activities–joint outcome is *collaboration (reciprocal peer tutoring)*. Collaborative learning has been defined as "jointly co-ordinated problem solving" (Crook 1994, p.49). It occurs when a group is given a "conceptual problem for which neither partner has sufficient understanding" (Saxe *et al.* 1993, p.119). Here the students are of similar cognitive ability and there is an equal interaction between them. If the discrepancy between the academic ability of the children were to be too great, it might lead to the less able accepting the knowledge without actively constructing the new knowledge necessary for internalisation. An example of collaborative learning would be children working together to determine how the software for laying out the pages in a class newspaper worked. This would be categorised on the right in the shared or distributed cognition part of the Dillenbourg *et al.* (1996) scale.

In joint activities–joint outcome group tasks the children can reconceptualise the tacit script – the existing rules for interacting (Engeström *et al.* 1997). In collective tasks, the more able children may play a larger role in influencing and interpreting both of these elements through discussion

and reflection (Daniels 2001b). In collaboration, each member of the group has an *equal* influence over the reconceptualisation of the script, the interactions, and the object (ibid).

2.2.3 Overview of group classifications

The type of group chosen by the teacher is dependent upon the task and the participants. It is not a sequential progression with groups starting as working groups, then becoming co-ordinated and progressing to collaborative or collective ones. Children working in collective and collaborative groups, however, use the skills required for co-operative group work, which in turn includes the skills for co-ordinated group work. This is represented pictorially in Figure 2.2.

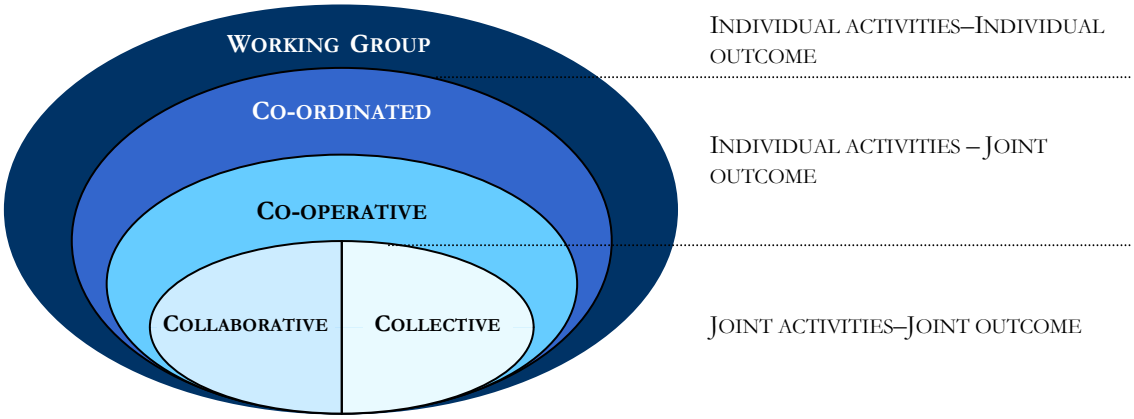


Figure 2.2: Pictorial representation of group work types in a classroom

An activity within a classroom can involve multiple types of group work. One example is the Key Stage 2 topic on determining factors that affect plant growth (DfES 2000). Initially the experiment may be co-ordinated. The class could be divided into groups and asked to look at a different area, for example, light and water, and then return their findings to the class. These groups may begin by being collaborative, with the children determining together how to set about the task and creating a shared hypotheses. Then the subgroup could become co-operative, each child taking responsibility for a task, e.g., measuring, or calculating. If they swap roles as individuals, and each teaches the next how to perform the task, or at the class level if they explain to the other groups, then they are working collectively. Finally, if the children are required to write an individual account, but are allowed to discuss this with the others, a working group is formed.

A training scheme to raise awareness of group skills so that future behaviour is influenced implies that the children can reconceptualise the group script. It can therefore be concluded from this section that the type of group activities that are suitable for the training scheme to be embedded within are joint activities–joint outcome ones, that is, collaborative and collective tasks.

2.3 Training schemes for group work

Many training schemes have been devised to assist children to work effectively together. If group skills are to be transferred between activities, the children must achieve the group's goal *and* be able to adapt the script from which they are working (Gallimore and Tharp 1990). As stated in Section 2.2 this limits the group activities into which the training scheme can be incorporated to collective and collaborative tasks. From this previous section, these types of groups share three attributes:

- *Communication* – as the group members exchange information, ask questions etc.
- *Reflection* – which is necessary for understanding and the internalisation of the knowledge
- *Responsibility* – both for their own work and that of the other group members

These are closely related but not identical. Communicating is not automatically reflecting, and reflection is not automatically communicated. Taking responsibility requires reflection and communication. To ensure his own understanding, and that of other group members, a child may ask if necessary, or answer or volunteer the information. The latter requires reflecting on what is appropriate. For help to be beneficial the child needs to ensure that it is timely, relevant, and correct; merely giving an answer without enough information for the receiver to amend his own misconceptions is detrimental (Webb *et al.* 2001). Once the child has decided the appropriate question to ask, or assistance to give, he requires communication skills to express the information so that the receiver understands.

In all three attributes, language, whether it is internal or external, is used as a mediating tool. It enables the group to express their own understanding, make sense of the guidance of others, and construct new knowledge. Vygotsky emphasises the role of language: "Children solve practical tasks with the help of their speech, as well as their eyes and hands" (1978, p.26). Language is also a cultural tool. Through language, written or spoken, experiences and tasks can be shared and understood. Language represents meaning, but by articulation it assists in the production of

meaning (Smagorinsky 1998). Effective group work requires a degree of linguistic mastery, for example, distinguishing between "may" and "could" (Stone 1993, p163), as well as social skills, such as listening (Underwood 1998, Mercer and Wegerif 1999).

As well as linguistic ability effective communication, reflection and taking responsibility require *metacognitive* thinking. This is the ability to reason about actions, strategies, the allocation of resources and roles, to monitor and predict outcomes, and to understand and act upon consequences (Wild 1996). This skill emerges around the age of 10, when children are capable of role negotiation, talking about what they are thinking, and have acquired the appropriate language skills (Azmitia 1998, Sharples 1999). At this age, children are capable of the joint creation and revision of knowledge necessary to reconceptualise scripts for behaviour (Azmitia 1998, p.218). Younger children can share resources and take turns, for example, but they are often unable to explain their views, repair breakdowns, and frequently dominate or block the problem solving process (ibid). Children of this age with training, however, can tutor others in specific tasks (Wood *et al.* 1995).

2.3.1 Communication

Communication within a group need not be verbal: gestures, pauses, eye gaze all have an impact, as in learning through observation. Language however, is the most obvious mediating tool. As stated, it allows the sharing and exchange of ideas and assumptions, although not all dialogues lead to cognitive development. Kruger (1993) says the talk must be *transactive*, that is, having criticisms, explanations, justifications, clarifications and elaborations. Mercer (1994) has an extended definition. He describes *exploratory* talk as having the same explicit reasoning, but emphasises the importance of participation and the shared goal of the activity. This talk can be identified through the examples of the participants building on each other's reasoning. There are more occurrences of the words "if" and "because", and more questions and longer statements (Wegerif 1996b). Through talk, the group members challenge and transform the thinking of others rather than paraphrasing their ideas. The dialogues also require social skills, e.g., participants must listen, support and help each other (Mercer 1995a). If a child uses exploratory talk and is aware of when to offer support and help he will perform better when tested individually (Webb 1991, Hoyles *et al.* 1993, Wegerif and Dawes 1998). These communication skills, defined by Tann (1990) as "critical listening" and "active contributing", are recognised in the English National Curriculum (DfES 2000).

Communication can be encouraged by:

- Providing the group with limited equipment or by distributing relevant information between group members. This forces interaction between the participants. An example would be Jigsaw activities, where groups must explain their findings to other groups (Aronson 2000).

- Introducing rules for behaviour within the classroom, including teaching the participants to listen without interrupting, how to respond to another group member, confirming understanding before continuing, etc. This technique is called training in *behaviour-based interactions*. Researchers who advocate behaviour-based interactions include Farivar and Webb (1994) who found that developing such "rules" for behaviour through discussions and teambuilding tasks resulted in more exchanges and a higher quality of work. Ashman and Gillies (1997) also found a similar impact of training. Children aged 10 who were taught interpersonal and collaborative skills were more motivated, autonomous, reached group decisions, and were less likely to ask for help. Interpersonal skills include listening, stating ideas, providing constructive criticism and accepting responsibility. Collaborative skills involve ensuring turn taking, sharing tasks, and resolving problems democratically. More recently the Talk, Reasoning and Computer (TRAC) lessons and the Talk Lessons developed by the Open University have had a similar positive impact (e.g., Wegerif and Dawes 1998). These rules are more effective when developed with the children (NCET 1990, Farivar and Webb 1994, Wegerif and Dawes 1998). Sharing this responsibility overcomes the teacher's assumption that such ground rules are evident to the pupils (Edwards and Mercer 1987). This training can be done in lessons *before* the activity (NCET 1990, Mercer *et al.* 1999). Training 8 to 10 year olds in exploratory talk however, does not always result in its use (Grove and Williams 1998).

- *Scripted co-operation* is the assignment and swapping of defined roles in a structured task during the activity. Johnson *et al.* (1990) suggest the roles of readers, checkers and encouragers. Burton *et al.* (1997) extend this to executor, reflector, constructor, checker and repairer. Goodman *et al.* (1998) have a similar set of attributes, describing: executives, instructors, record keepers, conciliators and critics. A similar idea can be used for computer-based discussions, with roles including constructive challenging, task management, encouraging participation, etc. (Pilkington 2002). Each role is associated with a separate cognitive skill and

type of behaviour. Initially the roles are discussed, allowing the group to develop a common schema. By taking and staying with their role for at least one turn, the participants can focus on that specific skill. The role provides structure. The children must understand what these roles mean, and how to communicate in that role; for example, being a leader is more than reading the instructions (Biott and Easen 1994b).

- *Scripted pair learning* or *procedural prompting* is more commonly used to improve task understanding. The teacher models appropriate behaviour and then shifts control to the students who must manage the exchanges *during* the activity (Palincsar *et al.* 1993, King *et al.* 1998). Content is structured through questions with generic stems, such as, "can you justify….", "what do you think will happen next…" (e.g., Palincsar and Brown 1988, Palincsar *et al.* 1993, Slavin 1995, King *et al.* 1998). Prompts require external speech to express explanations and planning. As with self-explanations, this has been shown to lead to a better understanding (Chi *et al.* 1989, Ericsson and Simon 1998). It is effective, because unlike "blind training", the students are taught to understand why such questions are useful (Lin 2001).

The three latter approaches are metacognitive strategies for learning. They encourage the children to explicitly consider and regulate their own thinking, regardless of the task and the domain. In effect, the tacit script is being made explicit and so can be amended. These strategies are artefacts that mediate the interaction. They can be physical, as in a shared piece of equipment, or abstract, as in a set of instructions on the wall, or a combination of both. An example of this would be specifying that only the person with the conch is allowed to speak (Golding 1997).

2.3.2 Responsibility

There are two types of responsibility when completing a joint activity–joint outcome (collaborative or collective) group task (Johnson *et al.* 1990, Slavin 1995):

- Individual accountability - individually ensuring that the skill is mastered
- Ensuring that the rest of the group is equally competent

These requirements mean that all group members must participate and agree any decisions. They must also share and maintain an understanding of their goal and their proposed solutions whilst helping and encouraging their peers, or asking for help and encouragement. Sharing responsibility for the success of the group and their friends motivates students (Garvin *et al.* 1995).

One strategy to ensure that responsibility is engendered is to structure the task so that the individual's score is dependent on that of the group. In the United States, teaching schemes based on these principles, such as CIRC Jigsaw activities, and STAD are used in 79% of elementary schools (Slavin 1996). In CIRC, students work in mixed ability teams on tasks such as: partner reading, making predictions, identification of characters, problems and problem solutions, summarising, learning vocabulary and spellings, reading comprehension exercises, and story-related writing. Team rewards are given based on the average of the children's performances (Slavin 1995). The STAD scheme works on a similar principle. The students are given a class presentation, then complete worksheets together, and at the end take individual tests. The scores are weighted on the basis of previous performance and rewards are distributed depending on progress (ibid). Jigsaw is a collective teaching strategy. Groups are instructed to master one aspect of a topic. They then teach the rest of the group what they have learned, which improves communication skills, and finally take an individual test covering all the aspects (Aronson 2000).

The reward scheme can be refined. Slavin (Slavin 1991, 1995, Stevens and Slavin 1995) and Johnson *et al.* (1990) have found that using rewards based on marks for both the individual contribution *and* the final product provides a greater incentive. The children must focus on their own understanding and that of their peers to maximise their marks. This approach has led to a reduction in "freeloaders" (students who do not participate but take credit for the work) as well as improving individual performance. Extrinsic rewards, however, do not always result in deeper understanding (Saxe *et al.* 1993).

2.3.3 Reflection

Reflection is necessary for internalisation and knowledge construction (Anderson *et al.* 2000). By reflecting on information available, it is possible to develop ideas, predictions, and evaluations. Social interactions also require monitoring. Although the group is accountable, the members

ensure that the outcome occurs through a process of reflection. They must decide whether the goal is agreed and this agreement is maintained, that each is contributing and that the communication and responsibility is appropriately shared. Reflection occurs through a dialogue in which people construct their own interpretation of events (Wardekker 1998). This can be an internal (virtual) dialogue, or an external one. Säljö (1998) believes that some discussions between children are equivalent to an internal dialogue. This type of exchange has been termed a social mode of thinking (Mercer 1995b).

The process of reflection impacts on future behaviour. Schön suggests that adults improve their professional activity by using reflection-in-action and reflection-on-action (1983, 1987). Reflection-in-action requires reflecting on the task as it is being performed. The practitioner's reflections influence her current behaviour in order to generate a new understanding and change in the situation *during* the activity. An example would be a teacher assessing the impact of her explanation on the class. If she realises some children do not comprehend, she can immediately adapt her explanation (Schön 1987). Reflection-on-action is a formative assessment. By considering the strengths and weaknesses of the recent performance, guidelines can be formed to improve execution in future similar activities (Schön 1983). This contrasts with a summative assessment, which compares performance against a standard or the performance of others. The reflection makes covert, abstract processes visible, which means it is easier to refine them, although it is possible to become self-absorbed and misinterpret one's behaviour (Johnson 1995).

Reflection can be encouraged at the *start* of the task, forming part of behaviour-based interactions, *during* the task, and on task *completion*. In this final this stage, children are asked to consider the strengths, weaknesses, and areas for improvement in their social and academic performance. This can be in the form of a group discussion (e.g., Tann 1990, Johnson *et al.* 1990, Galton and Williamson 1994, pp.167-169), or by the completion of individual questionnaires. This latter technique has been used with primary school children. The questions refer to responsibility, reflection and communication. They ask the children to evaluate how well they participated, helped each other, listened, complemented, asked questions, tried to answer questions, and paid attention, etc. (Kagan 1992, Lee *et al.* 1985 cited by Bennett and Dunne 1992e). More recently, the Talk Lessons from the Open University required children to go through a checklist at the end of the task to consider how well they worked together, and to identify possible improvements (Dawes *et al.* 2000).

2.3.4 Training schemes

The strategies for effective group work emphasise the attributes of reflection, responsibility and communication. The goal is to formalise everyday skills into scientific or scholarly ones by making the type of behaviour explicit (Kozulin 1986). They allow the participants to rationalise about how specific behaviour affects outcomes. Once achieved this understanding can be applied in other circumstances. To assist in this internalisation process the training scheme should be incorporated into multiple domains and used repeatedly (Dawes *et al.* 2000, Adey and Shayer 1994). Moreover, as in the development of any skill, not only should the scheme be used repeatedly, but the participants need to share an understanding of the point and purpose of an activity (Mercer 1995a).

The training strategies discussed in Section 2.3 aim to be self-regulated. Once self-regulation is achieved, the teacher can spend less time reiterating rules for appropriate behaviour, and the children are able to work together without constantly referring to the teacher (Bennett and Dunne 1992d, Ashman and Gillies 1997, Finlayson and Cook 1998). The teachers still have an important role despite the group being more autonomous. Teachers have to mediate where necessary and act as final arbitrators (Johnson *et al.* 1990, Crook 1994).

2.4 Factors influencing group work

Training schemes are mediating artefacts. Through the artefact, or the involvement of the artefact, a medium or middle ground is provided in which students can interact and test their ideas. In this research, it is proposed that a computer implementation of the training scheme would mediate by providing an environment within which the children can develop an understanding of group skills. The choice of artefact affects the nature of the activity, for example, a child's recollections of a day out are influenced by whether they are drawing or writing about the activity (Kress 2002). Even the same artefact can produce different outcomes. Mercer and Fisher (1992) suggest that identical software used by different combinations of teachers and pupils on different occasions generated distinctive activities over different timescales, with different problems, and hence learning outcomes. The impact of the mediating artefact, such as a computer system, is determined by the individual mediating their objectives through the technology *and* the existing tools and the community structure. This interrelation of factors can be represented by considering education as a specific activity setting (Engeström 1999a). This is shown in Figure 2.3.

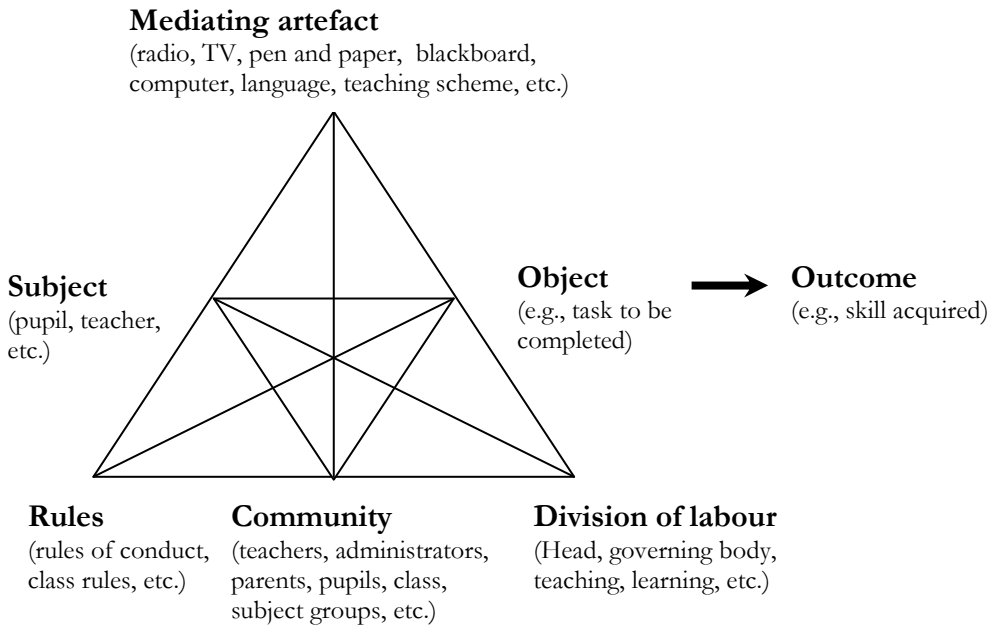


Figure 2.3: Application of activity analysis to education based on Bellamy (1996, p.126)

This diagram illustrates that the outcome is mediated by an artefact *and* its relationships with the community, the rules, the division of labour and the object itself. An example is the work of Light and Littleton. They found the gender of the subject influences the response to a computer puzzle (1999c). As well as influencing how a mediating artefact is treated, these aspects are interrelated themselves. The subject's exchanges with the community, for example, is dependent upon the friendship, gender, ethnicity, size of community, the task, whether they are observed, whether the teacher is working with them, and the ability of the participants (Pauli and Reusser 1997, Linehan and McCarthy 2001). Similarly, Daniels argues that the type of classroom, the rules, attitudes, and expected division of labour, impacts upon the use of mediating artefacts and the outcome (Daniels 2002).

This representation of an activity is not static. The objects, theoretical ideas and tools that shape the activity change over time (Engeström 1999a). The system is constantly being challenged and adapted as contradictions between the aspects arise. Engeström (1999c) calls this an expansive cycle. There is no pre-determined course of development since each participant has a different viewpoint and approach. Over time, a new social structure is formed on the basis of the preceding one. It requires a reflective analysis of the existing activity structure: "one must learn to know and understand what one wants to transcend" (Engeström 1999b, p.33).

One representation for each activity is insufficient. Not only does the representation change over time but it is dependent on the individual subjects. The subjects taking part in the same lesson, for example, could be trying to achieve different objects, i.e., they may have congruent goals. The teacher could expect her pupils to gain the domain knowledge being taught, however the children may wish to play around, talk to their friends, or avoid trouble (Edwards and Westgate 1994a). In addition, each person is involved in multiple activities occurring simultaneously that are distinguished only by their objects (Kuutti 1996, p.27).

An analysis of how these different factors interact to affect the outcome is the goal of current research into group work according to Dillenbourg and his colleagues (Dillenbourg *et al.* 1996). The first goal of group work research was to determine whether groups were beneficial and how effectiveness can be measured, e.g., by task performance or individual understanding (Ashman and Gillies 1997, Littleton *et al.* 1998, Anderson *et al.* 1999).

The second stage was to determine the impact of independent variables on the dependent value of performance. Researchers looked at the community and tried to determine the ideal size and composition of groups (Light and Glachan 1985, Hoyles *et al.* 1994, etc.). Other investigations included comparisons of homogeneous or heterogeneous ability groups (Tudge 1990, Slavin 1995, Slavin 1996, Webb 1985, 1991), the ratio of boys to girls in the group (Underwood *et al.* 1996, Joiner *et al.* 1998), age (Spavold 1989), previous experience (Kutnick 1994, Tharp 1993) and learning styles (Pillay 1998, Kolb *et al.* 1984).

From this work it was concluded the interactions between independent variables and the desired outcome influenced the group structure. Consequently, current research now focuses on the interrelationships within a group task. If a supportive environment is needed, for example, then friendship groups are more appropriate (Reid *et al.* 1989, p.10). If the task is problem solving, splitting high performing friends to form groups with less able pupils encourages higher order interactions (Kutnick 1994, p.26). Similarly, the division of labour showed that outcome is influenced by the relative status and expertise within the group (Rogoff 1990). A teacher-pupil dyad will have a different conversation due to the relative status than a pupil-pupil group (Biott and Easen 1994b), but pupil-pupil groups will be more engaged with the task (Cooper and Brna 2001). The former benefits skill acquisition, the latter conceptual change (Dillenbourg *et al.* 1996). Researchers have also considered how the object mediates the exchanges. Ideally, a group

should be given a task in which the object is to improve the situation, there is no correct solution, and it is not merely a problem or puzzle (Biott and Easen 1994c).

2.5 Using groups within the classroom

The previous section considers the factors that affect group work. The classroom is one of the major influences and is elaborated upon in this section. Schools can be considered as teaching two types of knowledge: i) *domain* – which are topics specified by the National Curriculum and can be directly assessed through the product; and ii) *general learning skills* – which includes non-domain specific skills, such as time management. This knowledge can be taught to individuals or groups, as shown in Table 2.1. The group skills that are the focus of this research program fall into the group general learning skills category.

	DOMAIN	**GENERAL LEARNING SKILLS**
INDIVIDUAL	Curriculum subject, e.g., mathematics, English, science	Information retrieval, interpreting, time management, research skills, etc.
GROUP	Group project work, e.g., book on animals of the world	Social skills, e.g., awareness of others, negotiation, leadership

Table 2.1: Types of knowledge gained within a classroom

Research indicates the benefits of group work for developing both types of knowledge (Kutnick 1994, Slavin 1995, 1996, Bennett and Dunne 1992a, etc.). Classroom observations in England, however, show that group work is used infrequently although the levels are improving. In 1976, 9.8% of the pupils' time was spent working in groups; by 1997 this had risen to 16.4% (Galton 1998). One explanation as to why teachers are reluctant to initiate pupil activities, including group work, is that the children may generate a great deal of noise. Teachers may assume that this means that they are displaying a lack of professional competence, as they appear not to be keeping order, despite the fact the children are being productive (Edwards and Westgate 1994a, p.47).

Another possible reason for this contrast is the emphasis on the Standard Assessment Tests (SATs). These are individual examinations in mathematics, English and science taken at the end of Key Stages 1 and 2, ages 7 and 11, towards which the majority of pupil and teacher time is geared (Adey and Shayer 1994, DfES 2000, Galton 1998, Henry 2001). Although aimed at providing information for teachers and Local Education Authorities and acting as a base line for future performance, these results are used to judge the school's performance and often influence

which secondary school a child will attend (DfEE 1999). Consequently, in some schools there is a focus on domain knowledge rather than on general learning skills.

Miettinen (1999) argues that this focus is detrimental. An emphasis on memorising and passive reception of texts may allow a student to pass examinations, but this does not assist them in other contexts. It also results in a didactic style of teaching; the teacher does the majority of the talking (Cummings 1988). This can be seen in the initiation-response-feedback (IRF) structure that frequently occurs, usually with teachers nominating the next speaker (Edwards and Westgate 1994b, Gallimore and Tharp 1990). When there is an interaction, teachers often ask rhetorical or pseudo-questions, which are closed or inappropriately pitched, and do not give sufficient time for a pupil to answer (Mercer 1995b, p.45). This type of questioning however, tests whether the relevant knowledge has been gained (Edwards and Westgate 1994b).

2.6 Evaluation of group skills

Section 2.5 highlights the minimal use of group work within the English primary school. Its use, however, is compulsory (DfES 2000), and Section 2.4 discussed research in this area showing that group work can be effective in developing social and cognitive skills. This section looks at the methods used to evaluate how children work together. It discusses how this information can be used to raise awareness of group skills. It also contrasts the methods that teachers use to evaluate group work, which is done in order to report whether the Attainment Targets are met (DfES 2000), with the approaches taken by researchers.

Evaluation can be used to generate feedback. Feedback differs from evaluation as it allows the highlighting and correction of misconceptions. It is one of the seven means of assisting performance identified by Tharp (1993). Evaluation in the form of summative assessments, whether the answer is correct or not, does not highlight *where* the error was made; it could be due to a mistake, curiosity, boredom, or fatigue (Knussen *et al.* 1991). In contrast, feedback informs change and suggests actions for improvement. It can also assist the teacher in providing appropriate future tasks, i.e., feedforward (Bennett and Dunne 1992b). The type of feedback influences the amount of reflection. In some activities for example, simplistic feedback, such as warm or cold, encourages discussion (Brandes and Wilensky 1980). While Galton and Williamson (1994) found extrinsic and corrective comments, i.e., "that's right" or "that's wrong", resulted in less discussion than when given critical comments about the quality of the work.

Requiring feedback to be interpreted in a discussion seems to facilitate problem solving and promote understanding (Wegerif 1996b).

Evaluation of discrete domain knowledge is easy to assess through tests, work or contribution to a whole class discussion (Kutnick and Rogers 1994). Evaluating group skills is problematic. This is unsurprising, given their subjective nature, the lack of formal guidelines from the National Curriculum, and that they are context dependent. Unlike a standalone test, which looks at the outcome, account must also be taken of the activity setting, that is, the community, division of labour, rules, mediating artefacts, *and* the product (Engeström 1999a).

Teachers are asked to consider a combination of factors to assess how well children work together (Ager 1998). Some conclusions can be inferred from the outcome of the group task. Use of group general learning skills can be distinguished by identifying if the work has been planned, is clearly constructed, completed within time, if all have contributed and are satisfied with the outcome and so on (Dunne and Bennett 1996). Some information can be gathered through observation of the group. This data needs to be treated with caution, however, since the close presence of an observer can change group dynamics and observation from a distance cannot insure that the interaction is on task (Biott and Easen 1994b). Other data includes the amount of assistance the teacher gave to the group (Ager 1998), and the use of interviews and post-tests, written or drawn depending on the group activity (Dunne and Bennett 1996). Self-assessment can be included in evaluating the group's performance if the children understand its purpose and grasp what is required. If given guidelines and involved in the record keeping and assessment process primary school children were found to make objective judgments on their own and others' task contributions (Finlayson and Cook 1998). These assessments are more precise if the children regulate others rather than themselves (Dillenbourg 1999). Little use is made of self-assessment in England, especially with respect to use of group skills. If it does occur, it is to evaluate domain knowledge gained and is achieved by children comparing the work with an existing piece of marked work (Weeden *et al.* 1999).

The teacher's approach to assessment contrasts with the more rigorous analysis of group skills by researchers. There is no set of measures that are specific to collaboration in cognitive science (Dillenbourg 1999, p.12); this results in the data gathered being dependent on the issue examined. This means researchers have included:

- The attitudes of the participants
- Styles of conversation
- Productivity (i.e., how much of the task was completed)
- The equality of their participation
- The roles played
- The amount of conflict and help
- Task achievements

(Wild and Braid 1996, Gillies and Ashman 1998, Mercer *et al.* 1999, Galton and Williamson 1994, pp.145-164, NCET 1990, Issroff 1999, pp.7-12, Webb 1997, etc.)

Discourse analysis, for example, is a popular approach used to determine participation, amount of conflict, roles, etc. The defined coding schemes reduce large volumes of data into manageable categories that can be analysed quantitatively. Categorising phrases can be inaccurate though, since they are context and content sensitive. If a transcript is used, it is harder to make judgments about the interactional tone of the encounter, e.g., posture, gesture, facial expression, eye contact and rapidity, volume, timbre and intonation of the speaker's voice (Edwards and Westgate 1994a, p.73). Transcripts may miss contributions of general agreement necessary for group cohesiveness (ibid). Phrases may be ambiguous, as interactions have a history leading to shared meanings that can be abbreviated, for example "like last week" (Edwards and Mercer 1987). Alternatively, phrases can have multiple meanings, or a meaning occurs over several utterances, the recording device may influence the exchange, or meanings change during speech (Wegerif and Mercer 1986, Draper and Anderson 1991). The categories themselves are dependent on the interests of the researchers. Gillies and Ashman (1998), who studied heterogeneous groups of 7 and 9 year olds, classified group types by student interactions, the number of interruptions, orders, and their cognitive processes, e.g., giving ideas and evidence or generalising. Wild and Braid (1996), divided task-specific cognitively-oriented talk as being: self-maintaining, directing, reporting, logical reasoning, predicting, projecting, and imagining. Wegerif

and Mercer (Wegerif 1996a, Wegerif *et al.* 1998, Mercer *et al.* 1999) have looked at the type of exchanges and categorised the talk as cumulative, exploratory and disputational.

Unlike pre-tests and post-tests, analysing talk can pick up examples of reciprocal peer tutoring and other discourse structures (Forman and McPhail 1993). Yet, although a conversation may appear to be productive, it may not change a child's understanding (Crook 1999), or represent the understanding of an individual within the group (Fisher 1993). An illustration of this is given by Howe (1991), whose analysis of the discussions on physics concepts indicated that the group member's understanding had regressed, yet the summative post-tests improved. This contradiction between quantitative and qualitative findings emphasises the difficulty in assessing group skills, and consequently determining appropriate feedback.

2.7 Computers within the classroom

The previous sections discuss the attributes, training and feedback required to develop group skills. In this section the role of computers as a tool to mediate between children is described. This is not a new concept. Light states: "It may be, then, that the computer can deliver something which other teaching and learning contexts tend to lack - perhaps by providing just sufficient structure, direction and support to the learning process to enable the collaborative learning process to be effectively sustained" (Light 1993, p.54). Children are often more confident and motivated to use technology than their teacher (McFarlane 1997a, Fisher 1993, Brown and Howlett 1994).

In addition to these advantages using computers is compulsory within the classroom (DfES 2000). They can be used by individuals, but are frequently used by groups. Until recently, limited resources and National Curriculum requirements forced children to share machines (Wills 2001, DfES 2000). This benefited researchers in their analysis of group work. Computer activities are usually well defined, with clear goals, fixed activity times, and they allow for the easy gathering of information (Light and Littleton 1999b, Wegerif *et al.* 1998, Crook 1994). In the remainder of this section studies of the use of computers in schools are examined, including how computers can be explicitly used to support groups.

2.7.1 Computers as a tool for the individual

Computers can be used as a tool for teaching the individual. "Drill and skill" packages found in most classrooms provide levels, a structure, and a meaningful environment for subjects with a clear domain (McFarlane 1997a). They reduce the cognitive load of the individual by externalising the problem, representing and constraining the information and options, and allowing for the easy manipulation of relationships and abstract material (Wild and Braid 1996, McLoughlin and Oliver 1998). They also provide a degree of autonomy, since they allow the student to control the rate of material presentation, repeat stages and skip forward, all of which is impossible in IRF lessons (Whitebread 1997).

This autonomy promotes *intrinsic* motivation, desiring to do well for one's own satisfaction. This can be enhanced by a positive performance evaluation from the computer (Ryan and Deci 2000). Such an evaluation can also enhance *extrinsic* motivation, i.e., performing an activity well to achieve a separate desired outcome. In this case, a high mark could lead to a feeling of pride, an expectation of approval by the teacher, or add a competitive element (ibid).

Additionally, such systems can be used as a diagnostic tool for individual task understanding. Computers can record interactions or the input sequence, providing a record of events that can be analysed. These log files may not relate to what has been learned though. Webb (1985) found that verbal interactions describing understanding between students were not always reflected in the test scores. Another disadvantage is that, although such systems require the child to act, they cannot ensure that the child reflects, and so construct the desired understanding (Sharples *et al.* 2002).

2.7.2 Computers as a tool for the group

The limited resources in the classroom often meant computers were shared (Wills 2001). This can be advantageous since computers can encourage discussion (Light 1993, Hoyles *et al.* 1991b, Wegerif 1996b, Mercer *et al.* 1999, Stephenson 1997, Whitebread 1997, etc.). Talking makes covert processes overt and is one of the necessary skills for establishing a successful group. Researchers have analysed the type of discussion resulting when children work synchronously as a group at a computer using a range of software. The software used in these studies range from adventure games and simulations, e.g., the Crystal Rainforest and Granny's Garden (NCET 1990, Crook 1994, pp.177-183), to open application packages (e.g., word processors or LOGO

(McFarlane 1997b)). Although all promote talk, there is more assessing and confirming of progress, asking, explaining and countering associated with open programs, compared to closed tasks. Here the focus is on the interface, informing and assessing exchanges (Anderson *et al.* 1993, 1999, Wild and Braid 1996). Yet open software or adventure games do not automatically lead to exploratory or transactive talk. Bennett and Dunne (1992d) and Klinger (1999) report that technical issues, rather than task content, can dominate conversation around a computer. Wild and Braid (1996) advocate training before the use of such software to avoid these lower order conversations on software usage, and encourage higher-level discussions about the task.

Apart from the ease in creating and observing activities for groups at a computer, their use has other advantages. Groups working in the absence of a teacher can practice scaffolding the problems for others, reciprocal teaching, resolving conflicts and being sensitive to the needs of the rest of the group, but at the same time be supported in the task (Kutnick 1994, p.21). Such groups are assisted by the provision of a joint problem space, a physical reference that can be used for the negotiation and clarification of terms (Crook 1994, p.151). As for individuals, in addition to this reduction in cognitive load, computer applications used by groups allows the children to control the rate of working (Wild and Braid 1996, McLoughlin and Oliver 1998).

Working in groups at a computer is not always beneficial. Younger children are more likely to focus on the game aspect of an activity. When testing the effectiveness of a system to teach probability Enyedy *et al.* (1997) found that the groups cheered for "their" team in the simulation instead of predicting the outcome. Another group of children that may not benefit are those with learning difficulties. They may prefer to work alone, as there is no competitive element and they believe a computer does not "judge" them (O'Shea and Self 1983). Finally, if practising domain knowledge, as when using "drill and skill" packages, it is preferable for the children to work alone. Despite pairs being instructed to take turns when practicing, those who worked alone were more likely to answer questions correctly and faster (Jackson *et al.* 2001).

2.7.3 Computers as support for group strategies

Section 2.7.2 showed that groups could benefit from using software designed for individuals. Applications specifically to support groups are found in a field of research called Computer Supported Collaborative Learning (CSCL). This focuses on how technology can enhance peer interaction and facilitate the sharing and distribution of knowledge and expertise among the participants (Lipponen 2002). CSCL applications have been developed for groups working in a face-to-face environment and in distributed settings. These require different approaches as each type of exchange has a different set of rules and behaviour due to the set of interactions possible (Straus 1997).

Some CSCL systems have tried to mimic group work by creating an electronic workmate for the child. Learning companions can collaborate and provide competition (Chan and Baskin 1990). A child benefits from such an application since to instruct his learning companion he needs to understand how to solve the problem himself. As when working with real people this develops social skills. The child has responsibility for assisting the "person", who can then question him and suggest approaches (Goodman *et al.* 1998, Dillenbourg and Self 1995). Unlike a teacher, or a "normal" machine prompt, using workmates has been shown to encourage the child to take the lead. The reflection process may become more difficult, however, when the child discovers that their partner is a machine (Boulay *et al.* 1999).

An alternative method for encouraging responsibility, communication and reflection is the use of dual-key control. This is not a novel concept, children already use game consoles with multi-player input (Stewart *et al.* 1999). Classrooms computers however, are assumed to have a single input channel, the mouse or keyboard, and limited output – usually visual although possibly auditory (ibid). Consequently, children often argue over who is to have control of the mouse and so "work" on the computer. This is overcome by Single Display Groupware (SDG) systems. These provide multiple mice or dual-key control, but only one monitor. This makes children jointly responsible for decisions, which improves motivation and task understanding (ibid). This was the approach taken by Light and his colleagues (Light *et al.* 1987). They found if 8 and 9 year olds *both* had to select a move for it to occur when solving the Tower of Hanoi problem, they would be better at solving the task alone when compared to children who had worked in unrestricted groups. The implicit incorporation of communication, responsibility and reflection by introducing dual-key control *without* specific focus on group skills has been shown to benefit

children aged 7 to 11. Multiple mice lead to more peer tutoring and teaching, higher motivation and more time spent on task (Stewart *et al.* 1999, Inkpen *et al.* 1999). Stewart and his colleagues suggest that explicitly negotiating a move, which is a group skill, means all group members participate (although the task may take longer). Incorporating dual-key control would assist in preventing situations with multiple mice where the children work in parallel and collaborate less (Stewart *et al.* 1999).

Technology can encourage communication through procedural prompting, that is, domain independent, generic or open questions such as "why do you think that?" (see Section 2.3.1). Wegerif believes such prompts cause children to "engage in uninhibited debate amongst themselves between the prompts and responses of a computer in a way which would not be possible with a directive teacher" (1996b, p.53). Pilkington (1998) found that the understanding of medical students improved when a system prompted the group with domain independent inquiry type questions. Anderson *et al.* (2000) found undergraduates who used a computer that acted as an electronic chairman constructed better projects. Despite the system having no domain knowledge, it could prompt for an exchange of views, pace the dialogue and request reflection. This contingent prompting is effective with younger pupils. Howe and Tolmie (1998) found that a domain independent system which asked for hypotheses to be created, and plans for experiments, resulted in more valid tests and appropriate conclusions when used by 9 to 14 year old children. Similarly the anthropomorphic agents used by White (White and Frederiksen 1998, White *et al.* 1999) prompt children aged 10 to examine what is necessary in each stage of the experiment. In these studies it can be concluded that it is the participants who co-construct knowledge or tutor each other.

2.7.4 CSCL feedback

The CSCL systems defined above have been shown to be effective through improved task performance when the children are tested afterwards (Dillenbourg and Self 1995, Stewart *et al.* 1999, Howe and Tolmie 1998, etc.). Additionally an analysis of conversation when using CSCL software shows children using more exploratory talk (Wegerif 1996b). These assessments on group skill usage are external to the software. Requiring the CSCL system itself to provide feedback on the interactions, rather than the domain knowledge acquired by the children, is more problematic. As discussed in Section 2.6 group skills are not a well-defined domain: ideal

solutions do not exist, understanding cannot be modelled, and libraries of common errors cannot be created.

Providing feedback on the effectiveness of interactions has been attempted in distributed CSCL systems. Soller and her colleagues (Soller 2001, Soller *et al.* 2002) have analysed exchanges that use sentence openers matching a speech act type: request, inform, motivate, maintenance, task, acknowledge, argue and mediate. The "speakers" and the number of each speech act they use are displayed. This allows the rest of the group to diagnose and repair the interactions, or the system can analyse these and suggest strategies to improve collaboration. The process of interpreting the feedback and amending future actions requires metacognitive skills by the participants. It also requires typing and the ability to express one's opinions in a written format using the correct sentence opener. When a similar approach was tried with children it was less successful. Groups of four children aged 10 to 11 misapplied sentence openers when communicating with other groups over a network, did not wait for the other groups to complete statements and ignored the prompts for coaching discussions (Robertson *et al.* 1998).

Constantino-Gonzalez and Suthers (2001a, 2001b) are also trying to develop group work skills as well as domain skills over a computer network with distributed participants. The feedback and suggestions for behaviour are given based on the amount of participation instead of the interaction content. If one member, for example, has not participated for a certain period and has a different individual answer to that displayed in the group's shared area, the system will encourage him to make his solution public. As before, this approach does not transfer well to face-to-face environments used by younger children. Instead of typing and exchanging information through the computer, children are more likely to express themselves using body language, for example, jabbing or pointing (Cummings 1988).

Providing feedback on group skill usage in face-to-face activities is more problematic, since it cannot be based on the contributions or task performance. Computers have no knowledge of whether one child dominated the task and the others acted as "freeloaders". Unlike teachers and researchers, computers cannot analyse and interpret conversations and interactions. This suggests more simplistic models for generating feedback are needed based on the available information. A computer cannot replace a teacher. It can, however, mediate between the children, or the children and their teacher, as they learn to analyse the effect of their interactions.

2.8 Summary

The ability to work together is an essential but not an instinctive skill. The use of groups is compulsory within the National Curriculum, and, in certain contexts, their use can improve the individual's social and cognitive skills. This chapter discussed the theoretical basis for the two hypotheses on the effectiveness of the proposed training scheme in raising awareness of group skills in 9 and 10 year olds. It was proposed that a computer implementation would lead to children performing the task more efficiently and formalising group skills so they could be used in other activities.

Groups can be viewed as a number of individual children, who learn by trying to overcome discrepancies between their understandings as in the socio-constructivist approach. They can also be considered within sociocultural practices. In these, development occurs by the child working with a more able peer or adult using psychological tools, such as language or cultural artefacts. Alternatively, cognitive development can be seen to occur within the social structure. The group is a single unit whose members co-construct knowledge. This is shared or distributed cognition. Groups based on these pedagogic theories are found within the primary school classroom. An analysis of the type of group activities suggests that for a training scheme to raise awareness of group skills so that future behaviour is influenced, the children must be able to reconceptualise the script for working together. This implies that the group activity in which the training scheme is embedded must be a joint activity–joint outcome task, i.e., it requires a collective or collaborative group.

Furthermore, for group work to be effective, the task must require the children to communicate, take responsibility, and reflect. Existing training strategies support at least one of these attributes. These strategies are more effective in raising awareness of group skills when used with children aged around 10 years and above. At this age, children have sufficient metacognitive ability with support to extract and formalise the requirements to work effectively as a group. To insure this internalisation of the understanding of group skills, any training scheme should be incorporated into multiple domains and used repeatedly. The schemes also require the teachers to mediate where necessary and act as final arbitrators.

Training schemes are only one factor that affects the awareness of children on group skills. Their knowledge is equally influenced by the activity setting, i.e., rules, community, division of labour,

the task, and themselves. Each of these aspects has a history that will also ensure that the outcome is not predictable. One of the major influencing factors is the classroom. Schools in England focus on the transmission of domain knowledge. These are curriculum subjects or skills, such as knowing about the Tudors or being able to write a scientific report. Group skills are part of general learning skills, individually these include time management, or researching information, as a group the skills include negotiating, turn taking and listening. The importance of this context suggests that any training scheme must be used within a classroom environment to ensure ecological validity.

The many factors that influence group behaviour also makes evaluation problematic. Currently teachers are asked to use a variety of methods to evaluate group skill usage as there are no clear criteria. They use the final product, observations, amount of assistance, and self-assessments, for example. Researchers are more rigorous. They are more likely to video or tape record and analyse the attitudes, style of conversation, the roles, as well as the outcome, and amount of conflict or help. Feedback based on the evaluation of group skill usage should influence future behaviour, however a method is required that uses data available. This is important if the medium giving the feedback is a computer as it is has restricted access to the exchanges. However, using a computer as a mediating artefact for the training scheme could be advantageous. With appropriate software computers promote discussion, even without domain knowledge. In turn, this requires reflection. Finally, responsibility can be incorporated by dual-key control.

To summarise:

- Training schemes that focus on communication, responsibility and reflection will result in improved cognitive and social skills
- Children aged around 10 have sufficient metacognitive abilities to reflect on their behaviour with support and consequently amend their future actions
- The context will have an impact upon the success of any training scheme
- Computers can act as a mediating artefact within a group setting

Chapter 3

PERCEPTIONS OF GROUP WORK WITHIN SCHOOLS

The literature review suggests that computer-based training schemes could be effective in raising awareness of group skills. Training schemes have been shown to improve group skill usage (Dawes *et al.* 2000, Bennett and Dunne 1992c, etc.). While computers have been shown to support groups of children aged 9 to 11 through encouraging discussion and the use of dual-key control even without explicit training schemes (Light *et al.* 1987, Stewart *et al.* 1999, etc.). Chapter 2 also identified the importance of context, that is, the community, rules, subject, outcome, division of labour, and the history of these as well as the mediating artefact. This indicates knowledge is needed about these factors to develop an effective training scheme. The National Curriculum, for example, sets Attainment Targets for group skill usage (DfES 2000). The interpretation of these is dependent on the classroom and task. There can be variations in the rules generated, what mediating artefacts are used, how the outcome of effective group work is measured, the expectations of the subjects and so on (Kress 2002, Mercer and Fisher 1992). Study 1, which is described in this chapter, investigates the context, such as, the expectations of the teachers, pupils, and the relationship between them (Edwards and Mercer 1987).

The goals of this study are to:

- Confirm that the skills of responsibility, reflection, and communication have been correctly identified
- Make recommendations for a training scheme to raise awareness of group skills that can be integrated into existing classroom practice
- Investigate whether a computer is a suitable medium for the training scheme, and if so, suggest design criteria

The study comprises of five interrelated parts: Studies 1a to 1e. They were performed over the course of 18 months and five schools participated. Study 1a combined a practical exercise to elicit children's attitudes to group work and provide an example of children working together. Study 1b comprised interviews with these children to determine their metacognitive ability, the impact that Study 1a had on their subsequent behaviour, and their beliefs about how group work is assessed. The interviews with teachers in Study 1c also covered assessment, and existing

teaching practices. Study 1d consisted of formal classroom observations, including the integration of computers into everyday teaching. Finally, Study 1e consisted of interviews with children from Study 1d to confirm the findings from Studies 1a and 1b concerning the purpose of group activities and how they should be structured.

As discussed in Section 1.2 the studies are designed to corroborate each other. In addition to Study 1e relating to Studies 1a and 1b, the attitudes of the children in Study 1b were compared to the observations of the teachers and researcher in Study 1a. Opinions of the teachers towards the teaching of group skills and use of computers are compared across the five schools in Studies 1c and 1d. They are also compared to the attitudes of their pupils in Studies 1a, 1b and 1e. Observations, interviews with teachers and pupils, and discourse analysis all generate an understanding of how groups are used, what aspects are viewed as important, and how they are currently taught and assessed (Hammersley and Atkinson 1995). These findings were corroborated by the descriptive activities (Cohen and Manion 1998, p.67). The children were given group activities in order to analyse the relationships among variables, how the vocabulary used by children relates to their actions, their own self-evaluation, etc.

The remainder of the chapter describes each of these studies in turn, and concludes by addressing the study goals: the correct identification of attributes for effective group work, recommendations for a training scheme to develop awareness of group skills, and the suitability of a computer implementation and possible design criteria.

3.1 Study 1a: Practical group work session

The aim of this study was to establish the children's opinions of, and actual performance in, group working. Data was collected from a group exercise in which children identified appropriate skills and activities for group work. The act of working together allowed the researcher to observe:

- The children's perception of group work
- Their behaviour in a group situation

Triangulation is incorporated into the study's design. The criteria recorded by the children and their own self-assessments were validated by the observed behaviour (Cohen and Manion 1998,

pp.233-238). The setting is described first to illustrate the lesson arrangements within the school and the possibility for group activities. The study also observed the impact of a session explicitly focusing on group skills on later activities.

3.1.1 Participants and setting

Seventy-one children and their teachers from three classes in different schools participated in Studies 1a, 1b and 1c: one private and one state Year 5 class, and one state Year 6 (see Table 3.1). In the state schools those from Year 5 were of a high academic ability, they averaged 288 out of 300 when taking the Year 6 Standard Assessment Tests (SATs) the following year, the National Average was 231. The children from Year 6 averaged 202, which is similar to that year's National Average of 218. The practical sessions were undertaken towards the end of the summer term, by which time the teachers and pupils were familiar with their class.

SCHOOL	PRIVATE - YEAR 5	STATE - YEAR 5	STATE - YEAR 6
Number of boys	12	18	10
Number of girls	6	8	17
Mean age *Standard deviation*	10 years 4 months *(3.65 months)*	10 years 9 months *(3.61 months)*	11 years 6 months *(3.86 months)*

Table 3.1: Participant details

All three schools were streamed by ability for mathematics and English, and there was one computer in each classroom. In the state schools, this was the only equipment available to the teachers for fulfilling the National Curriculum requirements.

The private Year 5 class was laid out with individual desks facing the front. The teacher allocated the seats. The National Curriculum was followed, but with a strong focus on teaching information related to the secondary school entrance examinations. In addition to computer activities in class, every week the children worked individually, or in pairs, for at least an hour in the separate computer suite with the dedicated Information and Communication Technology (ICT) teacher. The Year 5 state school had the classroom laid out in tables of four. The teacher had seated children in mixed gender and ability groups. The children sat on the carpeted area by the front desk for whole class teaching. The Year 6 class was divided into tables containing four to eight seats. The children had divided themselves into pairs, which the teacher had then allocated to tables. This resulted in groups of different genders and ability.

Before the studies a briefing had been given to each school about the purpose and activities it would involve (the letter can be found in Appendix B).

3.1.2 Method

In the first week, the whole class was observed for at least one two-hour activity. In the second, groups of up to six children selected by the teacher were taken into the library or an empty classroom and were recorded performing a group activity. Another observation period was conducted in the third week to determine whether the activity had caused a change in their behaviour when working in groups. To corroborate this observation the teachers were asked to report on the children's behaviour when working with others after the group activity (Appendix C1 contains the study plan).

The recorded activity in the middle session used a WISP (**W**ords, **I**deas, **S**tories, and **P**lans) board (see Figure 3.1). The WISP board is a form of mediating artefact; it acts as a prompt and focus for discussions. It is a focal point away from the researcher, and provides a tangible record of the children's thoughts. Originally, this format was used to help less able children improve their mathematical understanding by expressing a problem in a variety of forms (Jones and Haylock 1985). In this study, it was adapted to allow children to express their understanding of group work. The advantage of this is approach is that it allows data to be collected on the current understanding of group skills and the vocabulary used, but also allows the triangulation of data. The children's perception of appropriate group skills and their performance can be compared to observations of their actual behaviour during a group activity.

The board comprised of four quarters relating to group work. The children were asked to record their opinions in the relevant areas; there was no restriction on the order this information was added. The quadrants were:

- W*ords* - that describe group skills and why groups are necessary
- *Ideas* - these could be for subjects and tasks suitable for groups
- *Stories* - examples about the use of groups, and why they are appropriate or not
- *Plans* - these can relate to the layout of a classroom or activities that would be suitable for group work

There are no limitations on how this information was to be represented. The children were given a range of pens, plain and graph paper, post-its of various colours and sizes, and a tape recorder so they could write, draw or record their ideas and examples. Unlike standard questions that may appear to focus on tasks in academic environments, the groups were free to consider other contexts, for example, group activities performed at home or in the playground.

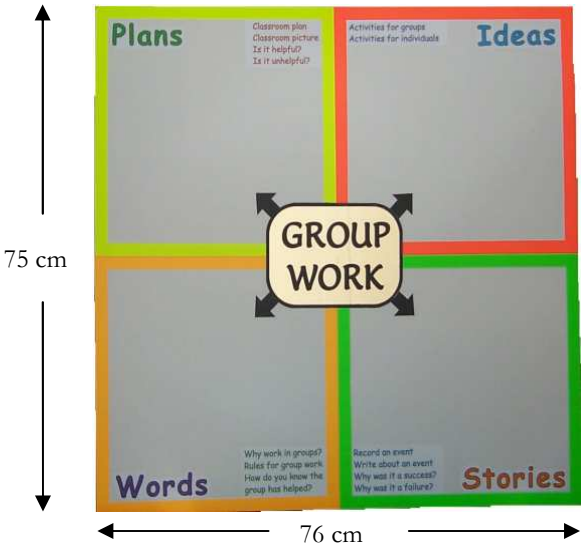


Figure 3.1a: The WISP *board*

Plans Classroom plan Classroom picture Is it helpful? Is it unhelpful?	**Ideas** Activities for groups Activities for individuals
Words Why work in groups? Rules for group work How do you know the group has helped?	**Stories** Record an event Write about an event Why was it a success? Why was it a failure?

b: The WISP *board prompts*

The exercise lasted 20-30 minutes. First, there was an introduction to the task. The children were told the University wished to know what group work meant and how it should be used in schools. This approach was used to render the problem as real, relevant and motivating, so that all group members would participate (Biott and Easen 1994c). Each of the prompt questions on the WISP board was discussed with the children before they were given the materials to complete the task. This open session lasted 12-20 minutes. To ensure activity was continuous, the researcher asked prompt questions such as: "How many children should be in a group?" and "Who should choose the groups?" The final section was a five-minute group discussion in which the children assessed and justified how well they thought they had worked together as a group (the lesson plan for the WISP board activity is in Appendix C2). Transcripts were made of each session.

During these three sessions, the following data was gathered:

- Quantitative data from the WISP board activity by collating the occurrences of spoken and recorded ideas
- Qualitative data from comparing the children's behaviour to the requirements they gave for effective group work
- Qualitative data from observations of group behaviour of the children in the following week

3.1.3 Results

Many of the children stated the exercise was "fun". The majority were articulate in describing the reasons for group work, how tasks should be structured and in what subject. Out of 36 reasons identified, over half of all comments related to:

- Sharing or getting ideas
- Improving communication skills through discussion
- Improving co-operative behaviour
- Gaining understanding
- Helping others to understand

Other reasons that were given included: helping in the development of friendships, achieving more work and that it was faster than working alone. Without giving any explicit reasons, pupils from all schools said that working as a group was easier. (A table giving the quantitative results gathered concerning reasons for group work can be found in Appendix C3.)

The majority of the observed behaviour of the children contradicted the reasons they gave. There were frequent squabbles over whose work should go where on the WISP board, or who had the idea first, and little effort was made to ensure that everyone participated. An analysis of exchanges with other children, that is, excluding muttering to themselves, showed statements were brief, averaging 8.35 words. Questions rarely generated justified answers. If asked where a post-it should be located, for example, the answer would be "Plans", although the post-its 1 and 4 in Figure 3.2 show that some elaborations occurred.

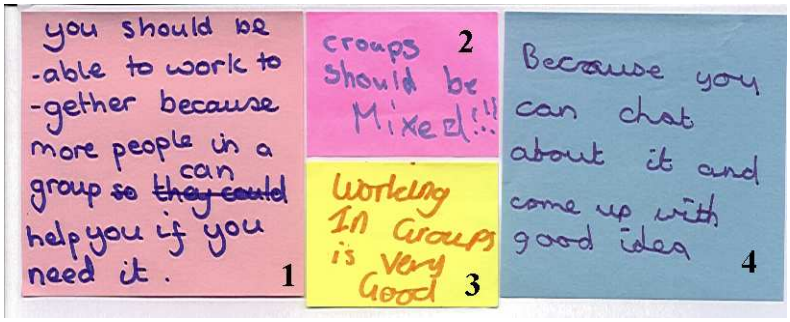


Figure 3.2: Examples of post-its, 1 and 3 from Ideas quadrant, 2 from Plans, and 4 from Words

An analysis of the transcripts showed that, on average, in each session there were 32 periods where children spoke simultaneously, and almost 34 occasions when one child interrupted the others. In the classroom, the teacher usually reprimanded the children interrupting.

Most children thought they had worked together when asked. Many said they had all participated and everyone had produced "lots of ideas". Pupils from the private school had made an average of 6.83 suggestions each, those from the Year 5 state school 5.31, and in the Year 6 state school 6.15. Only two groups felt that they could have been more productive, but offered no explanation as to how. Out of the 71 participants, only three children spontaneously said that this exercise was an example of group work.

Private school Year 5 pupils treated the activity as a *working group*, a shared task requiring individual products (individual activities–individual outcomes), although it was designed to encourage collaborative or collective working, having just one agreed product (joint activity–joint outcome). Instead of sharing ideas, the children in the private school produced written reasons but did not discuss them. In the state schools, especially the Year 6, more reasons were discussed but not recorded on the WISP board. The state Year 5 pupils also acted as if it was a working group activity, but demonstrated some co-ordinated behaviour, with explanations being directed towards the researcher instead of their peers. Those from Year 6 used a combination of group work types: working groups (when individually producing post-its), collaborative behaviour (by agreeing what they were meant to do), and collective working (e.g., asking the more able for explanations and spellings). These more able pupils tended to create co-operative groups, and took responsibility for the management of tasks, for example:

Pupil 1: Errr. Suzanne write on that.
Pupil 2: Write what?
Pupil 1: Urmmm… When you work in yeah yeah or…When you're working you co-operate.

The attitude towards, and perception of, group work reflected classroom practices. In the private school, the children competed to be top of the class. The same approach appeared in the WISP task through constant comparisons of the number of individually written and recorded statements. All the children were part of mixed groups in their school and suggested mixed groups in the exercise. The majority in the private school would then stream the groups, while the state school pupils preferred mixed ability. The Year 6 class was unique in suggesting that a group be created through self-selection *and* the teacher's allocation.

Rules given for ensuring appropriate behaviour in a group included raising hands before speaking, and numbering the participants then asking each in turn to ensure all contributed. Many came up with general guidelines such as "not being too loud" and "listening". No group discussed how to approach the exercise at the start and none of these rules were enforced during the WISP activity. Observations of later group activities by the teachers suggest the creation of these rules did not impact upon their behaviour.

When asked for subjects appropriate for groups 75% of the responses came from: computer lessons, science, English, physical education and mathematics. Computers were also frequently mentioned, though more by boys, as being part of the ideal classroom environment. In the Year 6 class this meant incorporating game consoles into the classroom to encourage group working. This suggests the WISP board succeeded in making the children consider contexts outside the classroom for group work. (The quantitative summary of subjects suitable for group work is found in Appendix C4.)

The children believed the size of the group is task dependent. The majority stated they should work in fours for general group tasks and in pairs at a computer (Appendix C5 contains the quantitative results gathered about group format). The most frequent justification for the latter was that it allowed each to be "doing summat", e.g., one would operate the mouse and the other the keyboard, or one would enter information while the other would check. Another advantage was that it increased the contact with the computer.

3.1.4 Discussion

These findings support claims children can: i) rationalise about group work; ii) reason effectively about the benefits; and iii) understand why they are asked to work together (Bennett and Dunne 1992e, Dawes *et al.* 2000, etc.). They show that children believe groups support them socially and academically. The majority of reasons for working in groups fall into the broad categories of improving communication and being responsible for the individual understanding of all group members. The children's opinion that they worked well together – which implies they believed they improved communication through discussion and gaining an understanding – contrasts with the observed interruptions and lack of elaboration. This implies the children have a rift in perception: either they use the same vocabulary differently to existing definitions of group work (e.g., Slavin 1995, Johnson *et al.* 1990), or there is a separation between their cognitive reasoning and their awareness of their actions.

The first argument suggests a discrepancy between a formal definition of group skills and the sense in which a child uses them. The contradiction between what the child says, and what is observed, implies that the child has not yet learnt to associate the forms, in this case words, with the meaning attached to them by others. Wertsch and Stone (1985, p.167) state: "it is possible for a child to produce seemingly appropriate communication behavior before recognizing all aspects of its significance as understood by more experienced members of the culture". Through exposure to the requests of their teacher children as young as six can give reasons for working together (Bennett and Dunne 1992e, p.139). This does not mean they have internalised the meaning and are able to apply the terminology outside the context in which it was developed.

The second argument is based on a similar difference in perception: the child has not yet internalised the fact that, to give an opinion on one's actions, one first has to reflect on what those actions are. Children will say the purpose of working in groups is to help each other, but still not appreciate that when actually working together they fail to assist each other. Johnson-Laird (1986) noted that children can grasp references, without having mastered the sense. If an incongruity is pointed out the children can act accordingly, but they cannot see the incongruity themselves or consider it out of context. This suggests self-monitoring by children, reflection-on-action, without support is problematic (Schön 1983). An inability to relate behaviour to actions is not unique to children. Elliott (1976) found teachers believed they were encouraging

equal participation, and only realised they were monopolising the conversation when listening to recordings.

The differing use of terminology compared to an adult and lack of mastery of the concept of group working would explain the small number of children who recognised the activity as being a group task. This is possibly because the activity was not introduced as a group exercise and did not belong to a defined curriculum subject. Similarly, the absence of understanding would explain the lack of transfer of group skills from the classroom to the WISP activity, for example, the increase in simultaneous speaking. The behaviour observed suggested that the task was predominately treated as a working group exercise. This implies children do not consider such an activity to be group work. Alternatively the children may have failed to recognise it was a group exercise due to the presence of the researcher affecting the group dynamics (Biott and Easen 1994c).

These results suggest any training scheme needs to incorporate an initial stage where children are asked to consider group skills before the task. This is to minimise the differences in perception between the group members of what a group skill entails. To encourage self-monitoring and reduce the number of incongruities, children should be told *before* the activity they will be asked to retrospectively assess how well they worked together. Stating the goal at the start of the task impacts on future behaviour (Lin 2001, Kramarski and Mevarech 1997, Salomon 1984).

The findings also suggest raising awareness of group skills needs to be a continuous activity. The observations suggest one period of reflection has little impact on future behaviour. This does not mean that such reflection is not beneficial. The expansive learning cycle implies internalisation happens over time (Engeström 1999b). This conclusion is supported by Adey and Shayer (1994, pp.60-61), who believe that intervention programs need not be intensive, but should be long lasting if they are to impact higher order thinking skills.

The results also support a central tenet of Activity Theory: that context plays a role on the perception of group work (Forman and Cazden 1985, etc.). This is inferred from the details the children gave concerning size, gender mix, ability, and selection choice of the groups compared to that observed in their class. It was also supported by the similarities between children's behaviour during the WISP activity and that observed in class. This may indicate that experience,

in this case their knowledge of appropriate classroom behaviour acquired over time, influences the outcome.

Children from two of the schools view technology as a way of encouraging working as groups. This findings support conclusions by other researchers, for example, Joiner *et al.* (1998), who found technology particularly motivating for boys. It is also interesting that work on computers is still seen very much as a practical task, where physical activity is necessary to being part of the group. This matches up with the traditional picture painted by Sheingold *et al.* (1984) of the "thinkist" and the "typist".

3.2 Study 1b: Interviews with children

This activity investigated:

- Why children felt that their teacher gave them group tasks – in contrast to the previous exercise which investigated why they thought groups were useful
- Whether the children believed reflecting on group skills influenced future behaviour
- How the children would assess group skills

The aim was to determine whether: children this age were capable of metacognitive thinking about group skills, what methods of assessment could be used by a computer system, and to corroborate the behaviour observed by the teachers and the researcher.

3.2.1 Method

Recorded group interviews were held with the same children after Study 1a. The teachers divided the children into eight mixed gender and ability groups with a mean size of 8.7 children. The interviews lasted approximately 20 minutes and took place in a separate classroom. The predefined questions related to the three areas given above. The researcher elaborated the questions when necessary. The results were analysed quantitatively according to the frequency of ideas and corroborated by a qualitative behavioural analysis. (The questions and a quantitative summary of responses can be found in Appendix D.) The results below summarise the main themes.

3.2.2 Results

1. Why group work is given

Many children had difficulty in deciding why group work was given. This was an abstract task, requiring a consideration of another's perspective. Some children said they had no opinion, they did not know, or thought group work was given "so we can sit round the tables". Of the answers given, the most popular reason remained that it enabled the sharing of ideas. All except one group suggested this. The change in emphasis of the question from their opinion to that of their teacher, however, did lead to a new reason. Half the groups in total, consisting of some state Year 5 pupils, all the Year 6 children and none from the private school, felt it was done so that the teacher could shift responsibility from themselves to their pupils. None of the children gave the reason because it was compulsory in the National Curriculum, or thought that it could provide a means of social support.

Within the groups there was very little questioning amongst themselves. No child asked why another held a certain opinion, for example. This contradicts their spoken opinions that group work is used to aid individual understanding, which implies an element of questioning and justification, i.e., exploratory talk (Mercer 1994). The amount of talk reflects that observed in the classroom. The Year 6 class was the most verbal; here interruptions and the loudest voice took precedence.

2. The impact of reflection on future group skill usage

To consider whether reflection had an impact the children had to think back to Study 1a and decide if they had acted differently when working in groups after that session. This approach was taken because the children had difficulty in understanding this notion on an abstract level. Most who gave an opinion said that thinking about group work would help people perform better in groups. Only one group gave an explanation: "as you would realise what you were trying to do". In this discussion on whether being given rules for group working would affect interactions, one group decided that rules would have no impact unless the group had agreed them first.

Some children claimed that their behaviour had improved due to the WISP board task, and that they had spent more time in discussion and tried to agree a course of action before taking it. This had not been observed by the teacher or the researcher. Other children stated that they had

not thought about group skills, either after group activities in class or because of the WISP task. This meant they did not know whether an exercise requiring reflection would be beneficial.

The children were asked how much group work they had done over the last week to determine whether awareness of using group skills had been raised. All the groups had difficulty in thinking of examples, and some children claimed that they had done none. The same task activities, working groups, were discounted as being group work despite: involving discussions on how to approach the task, sharing materials, and helping each other with spellings, etc. Only one child stated that he had worked in a group outside the classroom - he was building a go-kart with his family. Nobody mentioned the playground football games the children had categorised as a group activity in Study 1a.

3. How group work should be assessed

This topic on how group work should be assessed generated the longest and most animated discussions. All the children stated they had never considered this area. They concluded the teacher assessed group performance by: i) the completion of the task, ii) the task outcome, and iii) her observations. The last consisted of marks for distributing jobs, participation, and the amount of arguments – no distinction was made between destructive and constructive arguments. The children's major criterion for the evaluation of effective group work would be the individual's understanding of the task. Some groups also said that the work should be completed, and the time taken considered.

The majority of children believed that they had worked well together within the interview. This was judged using the level of participation, although several groups argued constructively whether a contribution that was an off task statement should be included.

3.2.3 Discussion

The results supports Study 1a by showing some children have advanced views of group working. These children appear to be aware that, for a group to be effective, they must take responsibility for the outcome, and reflect on and communicate their ideas. The idea of responsibility is context-related. Only the state schools generated this reason, and it was prevalent in the older age group. These attributes are not always observed in the group activities. Without reminders and an explicit structure, many children failed to follow their own arguments for group work

since they interrupted and spoke over the child trying to convey a point. As in Study 1a, reasons that are verbalised are not reflected in the actions. The children are either unaware of the accepted meanings of the vocabulary they use (Wertsch and Stone 1985), or understand but are unable to relate these definitions to their actions (Johnson-Laird 1986).

The fact that group work was not considered compulsory suggests its low importance to the children, despite the reasons given for its use. This view, taken with the belief that the teacher's use task completion and the quality of the outcome as the criteria for assessing group performance, suggests the children focused more on the domain skills to be gained rather than the group skills within the general learning skills category.

An absence of metacognitive abilities concerning the recognition of group skills is suggested by the number of children who were unable to:

- Consider reasons for group work from another perspective
- Answer the question concerning the impact of reflecting on group behaviour
- Recognise that group activities had occurred without the teacher's acknowledgement

Nevertheless, the spoken conclusion of some of the children that the period of reflection in Study 1a affected their behaviour in a positive fashion, suggests that reflection could be beneficial, but needs to be prompted.

As the children did not transfer behaviour from the classroom, such as taking turns to speak, to the group interviews, it could be concluded the children cannot transfer group skills to other curriculum domains. This, along with the previous suggestion that prompts are required, supports the research of Wegerif *et al.* (Wegerif 1996b, Mercer *et al.* 1999). Children can rationalise about group skills, but to ensure that ideas are expressed fully and that all participate any training scheme needs some sort of template for behaviour. Furthermore, from this and existing studies the children need a say in the template's creation (NCET 1990, Farivar and Webb 1994, Wegerif and Dawes 1998). Examples of successful templates for communication within groups are behaviour-based interactions, scripted pair learning and scripted co-operation (Palincsar and Brown 1988, Burton *et al.* 1997, Wegerif and Dawes 1998, etc.).

The third topic on possible methods of assessments supports the findings of Weeden *et al.* (1999). They found children were not used to self-assessment and required confirmation from the teacher. The results also support Schön (1983). He identified that children are less capable of reflection-on-action than adults. Despite this, there is evidence that children can use general learning skills, e.g., time management and taking part, when evaluating group skill usage. In the interview, the children stated that task understanding was the main criterion for assessment of effective group work. When having to evaluate themselves they considered their behaviour, as there was no discrete object or knowledge to assess at the end. This suggests that children are capable of self-assessing group skill usage, though they need support to distinguish between constructive and destructive arguments. Since general learning skills are domain independent, this implies assessment can be transferred between classroom activities.

3.3 Study 1c: Interviews with teachers

The purpose of the third study was to gather the perspectives of the teachers and examples about:

- Why group work was necessary
- How group work should be structured, developed and assessed in theory and practice
- Their attitude towards technology as a means to support group work

3.3.1 Method

Each of the three class teachers participated in a recorded interview following Tomlinson's (1989) procedures. This enabled the researcher to record the free associations made by the teacher as well as providing prompts for areas not covered by existing answers (see Appendix E for the interview agenda used). The interviews were conducted in private after Study 1b, and lasted between 15 and 25 minutes. The analysis was based on a qualitative approach, and corroborated by observed behaviour from Study 1a.

3.3.2 Results

All the teachers agreed that group working benefits the children academically and socially. Children achieve more and develop appropriate vocabularies for the task, and learn to care, be aware of, or support others. The main focus of lessons, however, is the completion of the domain content within the National Curriculum subjects to pass the SATs (DfES 2000) or entrance examinations.

Apart from mathematics and English, which are streamed, the teachers believed groups should be mixed by gender and ability. Groups can be pairs, tables, or whole class brainstorming – the group size is task dependent. The subject dictates whether the children or teacher selects the group members. The Year 6 class had most choice over whom they worked with. The Year 5 state teacher noted that friendship groupings tended to be between children of equal ability, and that high ability children were the most competitive.

Working groups were observed in all the schools. Children were set identical tasks, e.g., worksheets, and given time to confer with their neighbours about the subject and approach (individual activities–individual outcomes). No teacher, however, explicitly mentioned these tasks as an example of group work. Instead, the teacher of the Year 5 state class gave examples of co-ordinated and co-operative activities, e.g., preparing a holiday brochure and performing science experiments (individual activities-joint outcome). Co-operative tasks were also used in Year 6. This teacher mentioned the production of booklets in geography. She also incorporated collaborative work at the start of some lessons. This gave the children an opportunity to discuss their ideas about the task with their partner (joint activity–joint outcome). The result was to highlight misunderstandings often unobserved in a traditional "chalk and talk" approach. In her class collective group work, pairing able pupils with the less able, was also used occasionally. However, whilst being aware of the benefits of group work the Year 6 teacher decreased the amount before the SATs.

The teachers said they did not explicitly teach communication skills. They did, though, give guidelines in whole class brainstorming sessions, such as, reminding the pupils to listen to each other and give justifications. In co-operative tasks giving a shared mark based on the completed project was used to engender responsibility towards the group. Reflection in the Year 6 class involved the children agreeing the purpose of a task and whether it was beneficial. Reflection in

Year 5 was related to the curriculum subject. An example of this was the requirement to suggest ways of improving their performance in physical education and agree what aspects were successful. The state schoolteachers believed group skills were developed through the ethos of their schools. None of the teachers had heard of formal teaching schemes that developed group skills, for example, CIRC, STAD, Jigsaw activities, or Talk Lessons.

The teachers said they did not directly assess group skills. They argued their role was to prevent one child dominating or disrupting the others and to ensure that the majority of talk was on task. If they did give feedback it would be the result of continuous observation and consist of praise for working well, or questions about how the group performance could be improved.

The attitudes to technology were similar amongst the teachers. The state schoolteachers taught the ICT criteria laid out by the National Curriculum (DfES 2000). The private school covered this in separate lessons that incorporated the DfES (2001) teaching scheme. The ICT lessons tend to be structured so that pairs or individual children complete a practical skills exercise. It is seen as a "hands-on" activity – learning is achieved through doing. In addition to these lessons, the computer was observed being used for information retrieval in the private Year 5 and Year 6 class.

Within the state schools issues of reliability, lack of resources, and difficulty in running two activities simultaneously were mentioned as reasons for not using computers. The teachers did *not* perceive computers as a means of supporting group working. However, all three suggested a problem solving computer-based activity would support group work.

3.3.3 Discussion

There is a relationship between the attitudes and behaviour of the children, their teachers' opinion, and the style of work used within the classroom. The Year 6 class, for example, generated more verbal reasons, and their teacher integrated most talk into the lessons. Her class also believed that groups should be formed by self-selecting partnerships being paired with another duo selected by the teacher. While the two most popular subjects for group work according to children from the state Year 5 class (which are listed in Appendix C4) matched two of the subjects explicitly mentioned by their teacher. This supports previous studies that show context affects how a group activity is perceived and practised (e.g., Forman and Cazden 1985).

The interviews corroborate Section 2.5 and strengthen Galton's (1998) work. He found that the classroom is geared towards achieving the goals of the National Curriculum (DfES 2000), and training children for entrance examinations to secondary schools. This has meant that despite acknowledging the role of group work in providing academic and social support for children, the focus is on individual achievement in curriculum domain skills. This emphasis is highlighted by the minimal amount of teaching of group skills, even though speaking and listening skills are part of the Attainment Targets (DfES 2000). The focus on domain knowledge suggests that any training scheme to assist the development of group skills should integrate curriculum subjects. This would allow the scheme to be incorporated into current classroom practices.

This integration is confirmed by the teachers' attitudes. Discussions with the teachers show the National Curriculum strongly influences classroom practices. In contrast, schemes found effective by research appear to have little influence when imposed on classroom practises, as they are only adopted if the teacher feels ownership (Galton and Williamson 1994, p.170). The lack of influence of research findings on classroom practice is corroborated in this study by the teachers' lack of knowledge of teaching schemes to develop group skills and the underlying research. Figure 3.3 illustrates these relationships. The teachers consider the exchange of information between research institutes and the National Curriculum bi-directional. There is a theoretical basis for the curriculum, and the curriculum influences research. This belief implies that presenting the training scheme as validated by research would be less effective than presenting it as fulfilling the domain knowledge requirements of the National Curriculum. The latter means they do not have to restructure their approach.

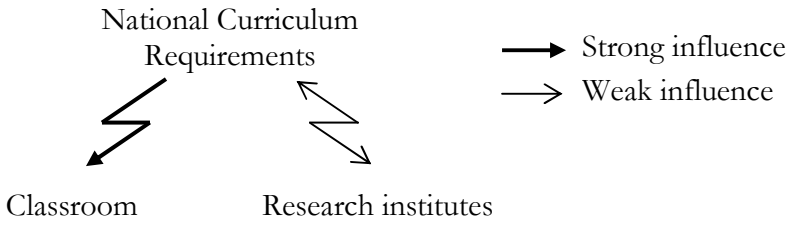


Figure 3.3: The relationship between schools, the government and research institutes

The lack of assessment and feedback on group skills could be attributed to the perceived higher status of SAT requirements within the National Curriculum (DfES 2000). It could also be influenced by the subjective nature of assessing group skill usage (see Section 2.6). Focusing on

the group's final product, the geography booklet or holiday brochure, allows a clearly defined set of marking criteria to be applied. Assessing group work by looking at behaviour is problematic given the changing interactions, and if observations are not continuous, the assessment may not always be accurate. Research has shown, for example, that talking off task is sometimes necessary to assist in the solution of the problem (Wild 1996).

Although only a small sample, the teachers could not recall any packages for groups and felt technology was unreliable. Apart from as an information source, computers were not integrated into other lessons. This suggests the 1997 finding that over a third of teachers are unconvinced or alienated by technology is still valid (BECTa and NAACE 1997). The teachers' opinion of computers as a set of skills to be gained may have been reinforced by the reiteration of the importance of knowing how to use a computer by the Minister of Learning and Technology (Wills 2001). This perception contrasts with the more positive attitude of their pupils found in Study 1a. The children accepted that computers are part of the education process and can be used as a tool for supporting group work - especially if they believe that they are in control. This implies that if a computer is used as the medium for a training scheme it should be reliable and easy to integrate into the classroom. Computers have another advantage; although not seen as a tool for group work, it is one of the few classroom subjects where children consistently work together.

3.4 Study 1d: Classroom observations

Studies 1a, 1b and 1c support Galton's (1998) findings that despite government requirements and research showing the benefits, group work is not a common practice in the English primary classroom. Study 1d used a combination of interviews and a Flanders Interaction Analysis Categories system (FIAC) to corroborate this. The FIAC enables the type of speech and activities used by the teacher to be quantified and supplemented the interviews and descriptive accounts.

3.4.1 Participants and setting

Two schools that had not participated in the previous studies took part in Studies 1d and 1e. Two Year 5 classes came from School X, a Church of England junior school in an affluent area of Birmingham, England. English is the second language for 12% of the pupils and 22% are from ethnic minorities. Two Year 5 classes came from School Y, a city junior school in

Birmingham. In this school, half the children come from ethnic minority backgrounds and have English as a second language. Three children participating from School Y were classed as having behavioural problems and were undergoing anger management training. The results of the nationwide Qualifications and Curriculum Authority (QCA) Year 5 examinations in 2001 show that the participants from School X were significantly better at spelling and mental arithmetic than those from School Y ($p>0.05$ and $p>0.02$ using t-tests respectively). They also scored more highly in the written mathematics and reading examinations. Overall, the scores of both schools are similar to the National Average (see Appendix F for a comparison of the formal assessments for the schools).

Both schools have dedicated computer suites with approximately 45 minutes timetabled use each week. All the children were familiar with computers, since ICT lessons were compulsory from Year 1.

3.4.2 Method

The researcher observed at least one lesson in each class before using the FIAC to categorise the events in a standard classroom lesson. Ideally, similar 45-minute lessons would have been observed but this was restricted by the availability of the teacher and the work to be covered. Observations of at least one 45-minute ICT lesson per class were recorded as field notes. The activities in the ICT lessons varied, and included word processing, using the Internet, or mathematical "drill and skill" packages. When making observations the researcher sat at the back of the class and did not participate in any activity. This information was supplemented by informal comments by the teachers after the lessons. An interview with both teachers from each school was conducted after the observation sessions. All the interviews were used to assist the researcher in developing an understanding of the context and clarify any events that occurred in the lessons.

The ICT observations are supplemented by teacher comments.

3.4.3 Results

Classroom context

This section is a summary of the general observation and interviews with the teachers. In both schools, all the lessons were taught as a class with the same teacher except for mathematics,

which was streamed, and music, which had an external teacher. In School Y, those children with poor English skills worked with a specialist teacher instead of participating in the literacy hour.

The teachers believed the majority of children in School X were reasonably mature but motivated by rewards. "Smileys" were given for good social behaviour or academic progress, 12 of these stamps could be exchanged for a certificate. Both teachers in School Y described the majority of their pupils as immature. One described a visit to a local secondary school for a mathematics lesson; the teacher there told her that her class's behaviour resembled that of five year olds rather than Year 5. Children from School Y have difficulty in what the teachers termed "thinking forward skills", i.e., not evaluating the consequences of their actions. A child would continue running when playing rounders for example, even though this caused their team-mate to be dismissed.

The children's perceived immaturity in School Y resulted in a high level of control and restricted amounts of group work. Children at the same table were given the same task and shared equipment in the literacy hour, art, geography and history. The science studies, such as growing plants in various conditions or researching the cycles of the moon were designed as co-ordinated or co-operative activities. They played in teams or pairs for sports and with a partner when at a computer.

The children at School X had a similar amount of organised group work. They also worked in pairs at a computer and played in teams or pairs for sport. In the classroom, they shared equipment by table or with a partner for geography, history, design and technology, and the literacy hour. Similar co-ordinated or co-operative groups were formed to School Y to complete the science curriculum requirements.

The teachers stated that they were responsible for all assessment and regulation of activities. As in the other schools, no credence was given to the children's self-assessment or self-regulation. There was no explicit teaching of group skills or assessment procedures, but the children were praised if they were working well together.

Observed classroom lessons

The FIAC system was used to record the verbal classroom climate. One class in School X was observed in a 45 minutes mathematics lesson on co-ordinates. This had an introduction describing how co-ordinates could be identified and used to plot shapes, followed by a worksheet. The other class were monitored preparing for an assembly on diseases in Tudor times. Due to interruptions and poor behaviour, the lesson was stopped after 35 minutes and the children then read silently. The results are shown in Table 3.2.

			Maths			**Assembly**		
			Occurrences	Occurrences as a percentage	Total per category of interaction	Occurrences	Occurrences as a percentage	Total per category of interaction
Teacher talk	Indirect influence	Accepts feeling						
		Praises or encourages	3	5.66%		4	7.69%	
		Accepts or uses ideas	1	1.89%			0.00%	
		Asks questions	7	13.21%	20.75%	6	11.54%	19.23%
	Direct influence	Lectures	9	16.98%		3	5.77%	
		Gives direction	11	20.75%		17	32.69%	
		Criticises or justifies authority	12	22.64%	60.38%	10	19.23%	57.69%
Student talk		Responds	6	11.32%		10	19.23%	
		Initiates	2	3.77%	14.09%	1	1.92%	21.15%
Silence		Silence or confusion	2	3.77%	3.77%	1	1.92%	1.92%
			53	100%	100%	52	100%	100%

Table 3.2: Summary of FIAC for School X

Two literacy lessons lasting 40 minutes were observed in School Y. They followed the same lesson plan on recognising verbs, adverbs, nouns and adjectives. The children were given a text and, after revising the grammar terms, the children identified examples of each in the passage. The results are shown in Table 3.3.

			Literacy: Class 1			Literacy: Class 2		
			Occurrences	Occurrences as a percentage	Total per category of interaction	Occurrences	Occurrences as a percentage	Total per category of interaction
Teacher talk	Indirect influence	Accepts feeling				1	1.49%	
		Praises or encourages				4	5.97%	
		Accepts or uses ideas	2	4.17%		1	1.49%	
		Asks questions	11	22.92%	27.08%	12	17.91%	26.87%
	Direct influence	Lectures	4	8.33%		2	2.99%	
		Gives direction	8	16.67%		21	31.34%	
		Criticises or justifies authority	9	18.75%	43.75%	13	19.40%	53.73%
Student talk		Responds	11	22.92%		8	11.94%	
		Initiates			22.92%			11.94%
Silence		Silence or confusion	3	6.25%	6.25%	5	7.46%	7.46%
			48	100%	100%	67	100%	100%

Table 3.3: Summary of FIAC for School Y

The teachers described the lessons recorded as typical. Table 3.2 and Table 3.3 show all the lessons were didactic. The teacher spoke for the majority of the time, usually having a direct influence through lecturing, giving instructions, criticising or justifying authority. Instructions were frequently repeated. More indirect influences, i.e., praise, incorporating ideas, or asking questions, occur in School Y. There they account for approximately 27% of interactions, while in School X indirect influences account for around 20% of interactions. In all the classes children have less input than the teachers. Children rarely initiate discussions in School X and never in School Y. Most contributions by children are answers to questions. These questions tended to be closed and testing domain knowledge, e.g., "give me one of the co-ordinates in figure 2" or "is dark an adjective or a noun?", or be a leading question, e.g., "can you remember what we were doing last week?" Silence occurred after directions had been given in both schools. This was short lived, and the teachers admonished the increasing volume levels when the children talked to each other again.

Observed ICT lessons

In School X the children were allowed to choose partners when working at a computer, resulting in four boy-girl pairs. Both School X teachers had difficulties making instructions heard, as the children did not immediately stop to listen. Once the lesson commenced, regardless of the activity, the children appeared on task. They did not leave their seats and only had the relevant applications open. The room was arranged so that the teacher could see all the monitors when she was at the front of the class. The majority of boy pairs and the more able girl pairs were competitive; they used any scores given and time taken to completion as a measurement of achievement.

In School Y the children also chose partners, but no mixed-sex groups were recorded. Both teachers described ICT lessons as "exercises in crowd control". They frequently shouted in an attempt to prevent children from crawling under tables, playing with the headphones, randomly printing, or leaving their seats to come and ask questions. Some boys took advantage of the fact the teachers could not view all screens simultaneously, even when at the front of the class, and switched between the task and wrestling Internet sites. When admonished they insisted that the sites "just came up". Threats to finish the lesson early, or to face the wall for the remainder of the lesson had little impact. The girls and more able boys tended to be the best behaved and least likely to be distracted.

Most children in School Y were competitive. The teachers believed the children measured performance by quantity rather than quality. In information gathering, for example, success related to the number of web pages found and printed - despite the teacher's constant requests to read and analyse the material gathered and only print pages that were relevant. Similarly, the teachers believed that their pupils thought completing a level on a mathematics package, even if the answers were wrong, was a greater achievement than attempting fewer questions and answering them correctly. Being able to use the Internet when the task was complete may also have motivated the children to work quickly. The only task-based conversations recorded between groups were to compare answers.

All the teachers believed computers were a set of skills explicitly required by the curriculum or a tool to support other subjects, especially mathematics.

3.4.4 Discussion

Although this is a limited study and conclusions drawn must be treated with caution it appears the children spend the majority of time working individually. This supports Galton (1998) and Studies 1a to 1c. The FIAC cannot be used to interpret the interactions, as this needs an understanding of shared meanings that are developed over time (Cohen and Manion 1998, p.28). It indicates, however, the style of teaching is predominantly whole class IRF, with pupils having little direct exchange with the teacher. The admonishment of noise corresponds to the finding by Edwards and Westgate (1994a); that teachers feel children being loud reflects poorly on them.

Few group activities are integrated into the classroom. Working groups with individual activities– individual outcomes are the most common type and result from limited equipment, rather than because of the findings on the effectiveness of talk in developing critical thinking skills (e.g., Mercer *et al.* 1999). Lessons focus on core curriculum subjects. Most exchanges observed were instructional, focusing on the transmission of knowledge, rather than regulative, which focuses on the principles of social order, relation and identity. The teachers were not observed, and did not mention, trying to raise awareness of group skills. There was no explicit teaching of communication, or encouragement of reflection to improve group performance; both schools lacked a collaborative culture (Daniels 2001b). The rewards used in School X, which, according to Slavin (1995, 1996), encourage responsibility, did not only relate to group performance.

Maturity is an issue. The development of metacognitive skills requires the ability to reason about behaviour and to be able to reflect on the most appropriate response. Many children in School Y seemed unable to do this without support. The lack of open questions suggests that the children are not used to providing justifications. This may indicate why there is little use of self-assessment or regulation, which both rely on providing reasons. Developing a training scheme that encourages the justification of self-assessment may enhance both skills.

Finally, the attitude towards computers indicates that teachers still see ICT as a set of explicit skills. It may be a suitable medium to raise awareness of group skills but it would need to incorporate curriculum domain subjects to ensure the application is used. The children, however, appear more positive towards computers and clearly enjoy these lessons. In contrast to the didactic practices in the classroom, the computer encourages talk although not necessarily exploratory.

3.5 Study 1e: Perceptions of group work among children

Study 1e investigated the reasons children believe group work is used, the subjects it occurs in, and how the children believe they should behave. It tested one of the hypotheses for children's difficulty in self-assessment suggested in Study 1a. That the children have not yet learnt to associate words, with the meaning attached to them by others (Wertsch and Stone 1985). Finally, like the previous studies, it looked at children's attitudes to computers and whether they would be a suitable medium with which to raise awareness of group skills.

3.5.1 Participants and setting

The participants were divided into 23 groups with an average size of four pupils per group. Fifty-one children from School X, and 46 from School Y took part. The children had been observed as part of Study 1d.

The activity was held in the library, or an otherwise empty classroom. The children had been in the same class for over a year and were familiar with their peers' behaviour.

3.5.2 Method

The children were informed the session was to identify what group work entailed. To make it meaningful it was explained this activity was part of a study for the University of Birmingham (Biott and Easen 1994c). A structured interview was used. Three questions were asked:

- How are groups used in school?
- Why work as groups?
- What rules would you give an alien who had just landed on earth for working well in a group?

The interviews were recorded and lasted 12-15 minutes.

3.5.3 Results

1. How groups are used in school

Pupils from both schools mentioned computers, physical education, geography, English, mathematics and science. School X also said design technology, and School Y art (a complete list can be found in Appendix G1 which gives all the subjects and group size in current classroom practice according to pupils).

The children were confused over the definition of a group. Seven of the 23 groups said they did mathematics in groups. Without prompting by the researcher, this led to debate on whether a group could be the whole class. Afterwards most children believed groups were less than 30, stating groups ranged from two to seven members – except in physical education where groups could consist of 11 children. The minimum size of two was also debated. Only nine groups spontaneously mentioned working at pairs in ICT as a group exercise. Others disagreed when pair work was suggested as a group activity: "it's not group work Miss - that's two people". There was also disagreement about the frequency of group activities. Despite having the same lessons some believed it happened at most twice a week, others that groups were used daily. The children were in at least a working group every day, since the literacy hour occurs daily and the majority, 17 groups, had mentioned English as a subject where they worked in groups.

2. Reasons for group work

Twenty-three clearly distinguished reasons for working in groups were given (Appendix G2 lists them and the frequency in full). Over 60% of the comments can be categorised as:

- Helping and receiving help from others
- Doing better work
- Making friends
- Being faster
- Sharing ideas
- Getting used to working with people
- Telling each other the answers

The children distinguish between providing help, and being given the answers. In addition, two groups in School X believed groups were used due to lack of equipment.

3. Rules for group work

Despite the inconsistency over what a group is and in which subjects they are used, the children were consistent about the rules for working well together. As in the reasons for group work, each rule is only counted once per group. (A complete list of rules and the frequency they were mentioned can be found in Appendix G3.) Just over 67% of the rules can be categorised as:

- Help each other or give advice
- Share ideas or thoughts or skills
- Be nice, e.g., be kind and do not tease or annoy or become angry
- Listen
- Talk together
- Work together
- Be friends, e.g., do not be bossy
- Co-operate
- Agree

These categories appear to be consistent. In a separate exercise six months later, the children recorded similar rules (Appendix G4 lists these rules and their frequency).

All the children were asked to explain these rules to an alien. Many children found this problematic and had difficulty verbalising an explanation. Definitions of co-operation, for example, included: "not not not just like … do ... something anything that you think of, just take it and just do that, but listen to other peoples comments 'cos quite often in group work there's like one person who's doing everything". Other children merely repeated the word, or gave a definition from another context. Co-operation was defined by one child as: "like say if your brain tells you to walk and that then your brain and your legs aren't co-operating then they're not walking."

3.5.4 Discussion

The children identified the same subjects as using group work between Study 1a and Study 1e. All five schools commonly use groups in: ICT, science, English, physical education and mathematics. This supports the idea that context influences the children's opinion. The subjects mentioned by the teacher in Study 1d correspond to the subjects mentioned by the children in their class. Similarly, they describe working groups as a group work exercise and so do the majority of their pupils, unlike teachers and their pupils in the previous studies. The idea that a computer could be an appropriate medium for raising awareness of group skills is supported since ICT is one of the few subjects that children acknowledge requires groups.

The study also suggests that children aged 9 and 10 are beginning to develop the linguistic skills needed to develop a vocabulary for defining group skills. Some can distinguish between giving help and giving the answers. The debate amongst the children about what comprises a group could have been due to linguistic interpretation though. Children have groups for mathematics but these are whole class groups. This terminology could have caused confusion when asked by the researcher to give subjects in which they work in groups. Even if pair work at a computer is not considered a group activity the range of subjects and sizes show children feel group size is dependent on subject, which corresponds to existing research (Light and Glachan 1985, Hoyles *et al.* 1994, etc.) and recommended teaching practices (Galton and Williamson 1994, Reid *et al.* 1989, etc.).

These discussions on what comprises a group appear to benefit the children's awareness. The talk revealed incongruities in understanding. The debate that followed seemed to cause the children to reflect and refine their own understanding. The agreement of the majority that school groups have less than 30 members, for example. Similarly, this process of externalisation forcing reflection is apparent when describing group skills. As in Study 1a, the terminology when giving reasons for working in groups is similar to that of adults and other studies of children (e.g., Bennett and Dunne 1992e, Mercer *et al.* 1999). When explanations were requested the children had difficulty, or could not define the terms. Distinguishing a term in the context of group activities was problematic, as shown by the difficulties defining "co-operation" - which was not always viewed as working towards the same goal. This supports the hypothesis from Study 1a, that children know the terminology – but not the meanings attached by others (Wertsch and Stone 1985). The implications for a training scheme are that:

- *To raise awareness of group skills the children must externalise what a group skill is* – externalisation requires the children to reflect on their understanding so it can be verbalised and possibly refined through discussion with their peers
- *A teacher is required to monitor occasionally* – this is so she can insure that inappropriate definitions that are being discussed, such as co-operation in groups being the ability to walk, are not internalised by the children (Vygotsky 1978)

3.6 Summary

The findings from these studies show that children across these five schools have similar attitudes to, and experiences of, group work. The children's perceptions are strongly influenced by their teacher. The beliefs as to why group work is done and the rules required correspond with existing studies. Children believe effective group work is dependent on communication skills and participation within the group. They do not use the terms "reflection" and "responsibility", but these attributes are implicit in the rules given in Study 1a and 1e. The children state group work is given to generate ideas and gain understanding, and that it requires co-operation and helping each other. The older pupils in Study 1b also recognise that groups help them to take responsibility. However, knowledge of appropriate behaviour does not always translate into actions.

Although teachers do not specifically teach group skills, they agree that group work improves social and cognitive development. When used, group work incorporates communication, responsibility and, depending on the task, reflection. There are no formal guidelines to assess group work. The teachers often use the product of the group activity since it is not subjective, and therefore easier to mark, and fulfils the perception of the National Curriculum as being a set of domain topics to cover.

Observations of a group activity in Study 1a show that the children believe they used more group skills than the researcher recorded. During the task, they constantly interrupted each other, talked over each other, failed to agree a plan and did not expand their ideas. One explanation for this difference is that the adult and child use the same terminology to describe group work skills, but assigned different meanings. This was corroborated in Study 1e, which showed children used appropriate words but had difficulty in giving an explanation. Alternatively, the child may understand the adult's terminology, but not have mastered the sense. The children appear not to have realised that to give an opinion on one's actions, one first has to reflect upon what those actions are. Despite the beginnings of linguistic mastery, Study 1d suggests that without support children have not the maturity or experience to work effectively as a group.

Studies 1b and 1e also show that children require regulation by others to adjust their own opinions, and prompts to justify what they claim. They recognise that agreeing rules for behaviour would assist performance, but cannot implement this guideline without support. This

could relate to previous experience, since all the teachers took responsibility for assessment and regulation. The children are used to focusing on the TGT and its outcome rather than their behaviour. This indicates that a system to support children develop group skills should explicitly request them to consider assessment criteria and behaviour throughout the activity. Study 1e showed this externalisation process highlighted incongruities in understanding and that discussions with peers led to a refinement in understanding. The teacher is required to monitor these discussions occasionally though, to ensure misconceptions are not internalised.

Study 1a suggests a child could not accurately assess group skill usage over an entire activity. It suggests that to raise an awareness of behaviour the children must consider their actions throughout the task. One method would be for the training scheme to incorporate prompts to consider behaviour *during* the task. Study 1a also shows one session in which group skills are focused on does not impact future group behaviour. The training scheme therefore needs to be used repeatedly.

Overall Study 1 shows using a computer system as a medium for the training scheme to raise awareness of group skills has disadvantages. Teachers and their pupils perceive computers as a set of explicit skills, a tool to gather information, or a method of practicing other domain skills such as mathematics. It is *not* seen as a tool for supporting group skills. All the studies, however, show that ICT is a topic in which children work together. Computers are motivating to children. Unlike classroom lessons there is less didactic teaching and the children have control over the pace of the activity. Computers could provide prompts to consider behaviour throughout the task. Studies 1a and 1c highlight the importance of a practical involvement with the task. This suggests the children should have an individual means of interacting with the computer. If a computer is to be the medium of the training scheme, the teachers suggest the software should be reliable and the group task should be a problem solving activity. In addition, any group task needs to support the explicit domain activities and skills required by the National Curriculum.

Chapter 4

DEVELOPING AND SUPPORTING GROUP SKILLS

As shown in Study 1 awareness of group skills in children from Years 5 and 6 is not inherent. The children often fail to communicate through speaking simultaneously or not listening. When giving answers or suggestions, they do not reflect upon the appropriateness of the wording or the timing of their responses. There appears to be little responsibility within the group. Children neglect to check that other members understand or agree with their opinions but focus instead on the task outcome. Despite the absence of these necessary attributes, children believe they work well together and overestimate the time spent using such skills. This suggests that the children need explicit support in developing awareness of their behaviour and understanding its appropriateness.

This chapter proposes a group task in which the children are to be supported in raising an awareness of group skills. This task consists of two activities running in parallel:

- The teacher-given task – the domain knowledge to master or product to be created by the group
- The group skills task – an awareness of their general learning skills

The training scheme described uses a model that scaffolds the latter task. The model combines findings from Chapter 3 and existing methods for developing group skills. The second section discusses whether a computer is a suitable medium for the model and discusses one possible implementation. In the final section, an experiment using this software is described. The goal was to test whether:

- Performing these two group tasks in parallel was detrimental to the teacher-given task performance
- Self-monitoring was possible when done concurrently with a group task
- Self-monitoring provided enough information for feedback to be generated on group skill usage

4.1 Theoretical basis

As discussed in Chapter 2, many teaching schemes exist to improve group skill usage (Bennett and Dunne 1992a, Johnson *et al.* 1990, etc.). They require the children to focus on at least one of the key attributes of responsibility, reflection and communication. This can be done implicitly, as in Jigsaw activities, where the children are responsible for one area of a task and communicating their findings to the remainder of the group (Aronson 2000). Or explicitly, as in Talk Lessons, where the children are encouraged to discuss appropriate group skills *before* the task and self-regulate their behaviour *during* the activity (Dawes *et al.* 2000). In this latter type of scheme, which requires metacognition, the teacher is effectively giving two tasks relating to the two types of knowledge gained within a classroom (see Section 2.5):

- The teacher-given task (TGT) is the group activity that contains the domain knowledge or skill the children are to master. Examples include investigating the factors that affect plant growth, or creating a class newspaper
- The group skills task (GST) focuses on general learning skills; the children consider their behaviour before, during and after performing the TGT

It is hypothesised the process of defining and monitoring their behaviour leads to the conversion of children's everyday knowledge of group skills into abstract scientific concepts (Vygotsky 1987, Kozulin 1986).

Teachers have been advised to split group activities into stages when trying to raise awareness of group skills. Bennett and Dunne break the GST into four stages. The teachers must: 1) raise awareness of appropriate behaviour *before* the task, 2) monitor the children's group skill usage *during* the task through observation and involvement, 3) evaluate the group processes, and 4) at the *end* provide feedback (Bennett and Dunne 1992d). This is illustrated in Figure 4.1.

.

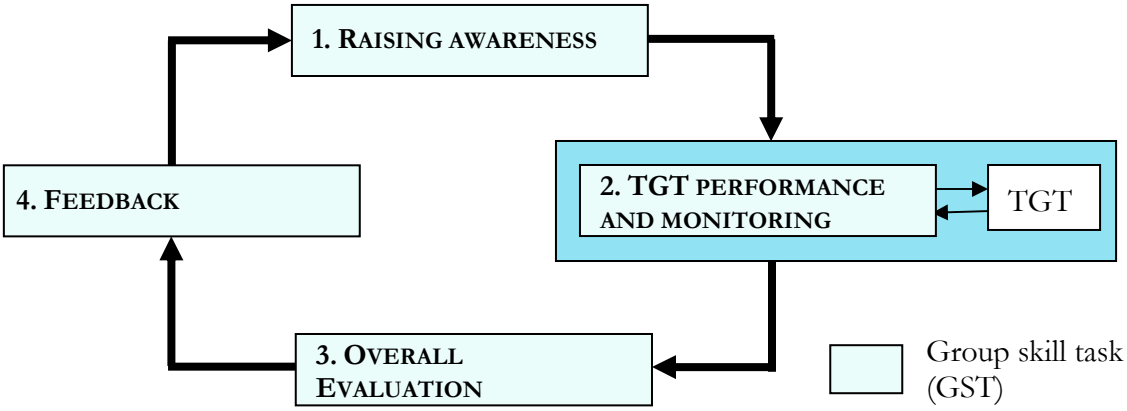


Figure 4.1: Cycle of training in group work (based on Bennett and Dunne 1992d, p.170)

The training scheme proposed is based on this structure. It incorporates the stages of explicitly raising awareness of group skills and continuous self-assessment from Chapter 3. Instead of the teacher externally regulating the group's behaviour, a model was developed to scaffold each stage with the group being supported by the teacher. The result is the Developing and Supporting Group Skills model (DSGS v1 model) shown in Figure 4.2.

The DSGS v1 model uses guidelines for assisting performance suggested by Tharp (1993). It provides *task structuring* by splitting the GST into four stages. This requires group skills be discussed before the TGT is started and ensures assessment occurs throughout and on the TGT completion. It also provides *feedback*, which is used as a basis for the next group activity. The model requires teacher support. It can only suggest actions and provide information, it cannot monitor whether these are acted upon or understood.

The model requires the children have some degree of metacognitive ability since it involves self-monitoring. From Study 1 it is hypothesised that children may have difficulties in monitoring their own behaviour as they: i) are use to focusing on the TGT, ii) use incorrect terminology, or iii) are unaware of the discrepancy between their reasoning and actions. The ability of the participants, or how the exchanges are structured is less important since only a limited number of exchange types can occur in a joint activity–joint outcome. In collaborative groups the children may use peer scaffolding to generate an answer together, or explore different opposed alternatives in argumentation. In a collective group, one child could tutor the other or act as a resource (Wood *et al.* 1976, Dillenbourg *et al.* 1996). If necessary, a combination of these exchanges could happen.

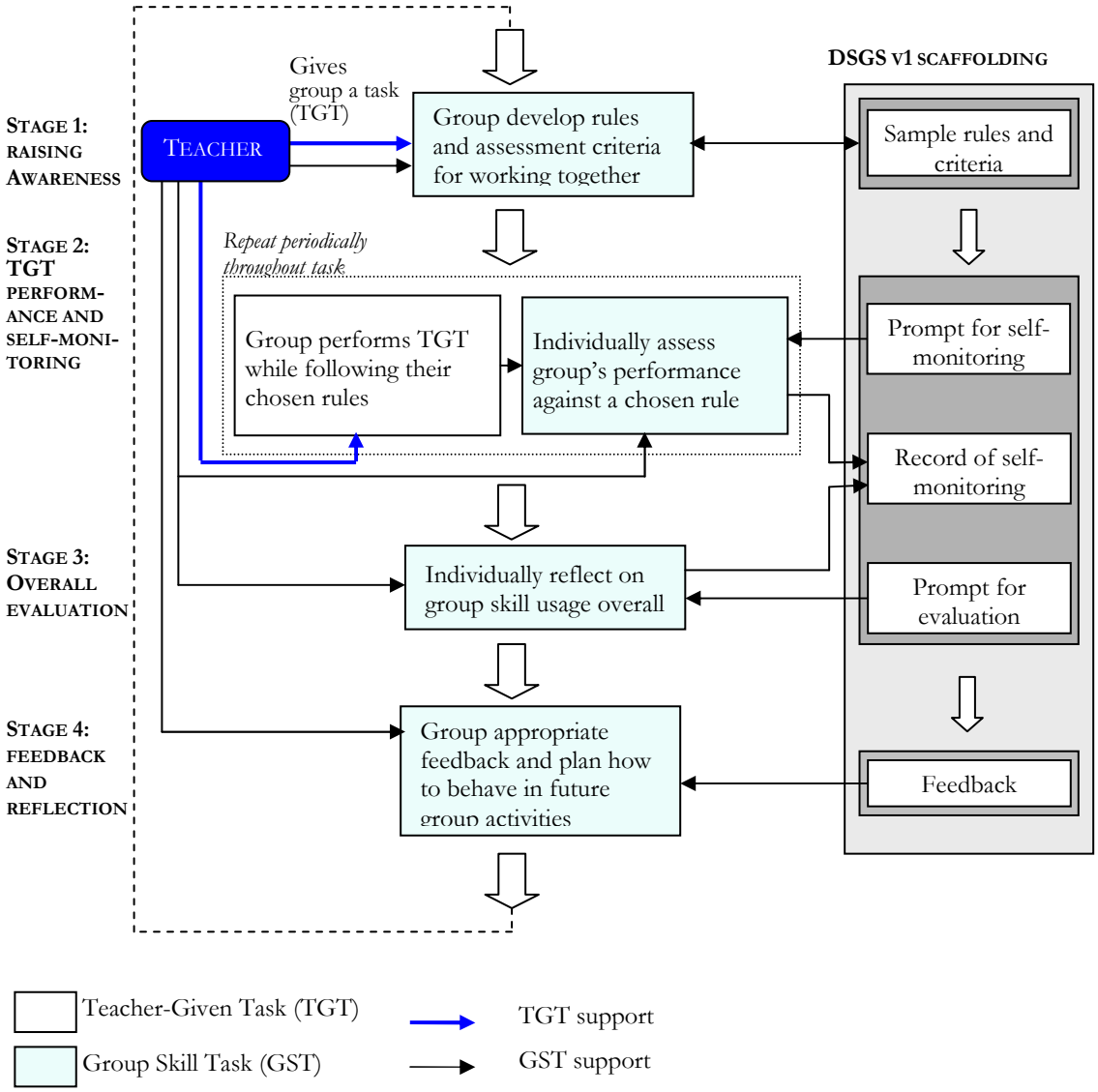


Figure 4.2: A group activity with scaffolding of the GST by the DSGS model

4.1.1 Stage 1: Raising awareness

As suggested by Lin (2001), Kramarski and Mevarech (1997) and Salomon (1984) to improve the children's GST performance they are told of the importance of working together and how it will benefit the TGT. Next the children are given sample "rules" for working together, these are associated with communication, reflection, and responsibility. Each rule is concrete, clear, and easy to state if followed, e.g., asking whether a child has listened or not. Ambiguous rules, such as "co-operate" suggested in Studies 1a and 1e were avoided. Some research indicates that only positive rules should be used (Galton and Williamson 1994, p.166). In this model however, the

suggestions contains a mix of positive and negative rules. This is to highlight that effective group work requires behaviour that should, and should not, be displayed.

Stage 1 requires the children to identify three rules necessary to work well together for this TGT. These can be taken from the sample rules or can be their own invention. Then are asked to talk as a group to ensure a common understanding of what each rule entails, and develop criteria so they can assess usage. If they believe listening is important, for example, then the children discuss how they would recognise they have listened; this could be through being quiet when others are speaking, or looking at the person talking. This discussion of rules before commencing a group task, a form of behaviour-based interaction, has been shown to improve interactions and the TGT performance (Farivar and Webb 1994, Ashman and Gillies 1997, Webb 1997, Dawes *et al.* 2000). This process is equivalent to rehearsing the ground rules (Mercer 1995b, p.113). Taking responsibility for choosing the rules gives a sense of ownership and may assist the children in following them (Study 1b, Mercer *et al.* 1999, etc.). By explicitly agreeing definitions of group skills and assessment criteria the children are more likely to develop their understanding (Wertsch and Stone 1985). The number of rules chosen is limited to ensure that the children are able to process other tasks at the same time (Brown 1987).

4.1.2 Stage 2: TGT performance and self-monitoring

One explanation for the overestimation of group skills displayed in Study 1a is the style of questioning that was used. The children were asked to assess the overall level of group work displayed without agreed clear criteria. Ager (1998) and Finlayson and Cook (1998) advocate the use of self-assessment in children aged 9 to 11 if the criteria are well defined. To simplify the assessment task in Stage 2, the children are asked to record when the group followed one of their chosen rules from Stage 1. They are asked to use the assessment criteria established and agree before making this observation. This recording can be done at any stage throughout the TGT. The assessment is of the group rather than the individual; Dillenbourg (1999) suggests children are more objective, and hence more accurate, if assessing others.

The DSGS v1 model acts as a mediator by displaying the agreed rules, providing a method for recording when the children believe they are following rules, and prompting them to reflect on their behaviour. The exchange between the TGT and GST is not one way. Behaviour in the TGT obviously influences what is recorded in the GST but it is proposed that the act of self-

monitoring in turn impacts upon the group's behaviour. The discussions should clarify the understanding of group skills and encourage the children to use these styles of talk and behaviour within the TGT. This monitoring, and consequent changing of behaviour during the task, can be compared to reflection-in-action (Schön 1987). Schön used this expression to describe how an adult practitioner would change her behaviour during the task according to her perception of events. Unlike an adult where the process is internal, in this model it can be externally prompted. Such prompts to reflect on understanding can impact the knowledge of the participants. Enyedy and colleagues (1997) found forcing a break in the TGT activity allowed a group of children to refine their task understanding. This impacted upon future task interactions. By pausing the TGT activity, the children can reflect and discuss with peers their immediate behaviour and may decide the skill is absent, and amend their future interactions.

4.1.3 Stage 3: Overall evaluation

Evaluation reflects the belief of the group about their overall behaviour. As in Stage 2, the group is responsible for the assessment. To ensure the opinion is shared, each group member is asked to justify his view. The rules from Stage 1 perform a similar role to the checklists used to assist children in identifying how well they worked together (Dawes *et al.* 2000, Kagan 1992, Lee *et al.* 1985 cited by Bennett and Dunne 1992e). The assessment should be more accurate than in Study 1a. In that study, it was hypothesised that children may fail to accurately assess group skill usage, as their descriptions do not relate their actions. In this model, incongruities between actions and reasoning are less likely, because the GST encourages the children to reflect on their behaviour *throughout* the TGT. To improve the accuracy of the assessment the children decide a group score rather than an individual one (Dillenbourg 1999).

4.1.4 Stage 4: Feedback

Feedback is required to assist the children refine their vocabulary for describing group skills, improve awareness and plan how future group tasks can be performed more efficiently. To encourage this reflection, the DSGS v1 model contrasts the self-monitoring from Stage 2 with the overall evaluation in Stage 3. The self-monitoring data recorded by the children should incorporate observations, simple discourse analysis, amount of conflict, etc. This addresses the issue raised in Section 2.6, that accuracy in evaluating group skills requires the assessment of a variety of data.

If in Stage 2 a group never acknowledged that they followed any of the rules the feedback would suggest group skill usage was poor. If this group then entered a high score in Stage 3, the marks would contrast. When asked to explain the discrepancy the children must reconsider their understanding of group skills and recall of their interactions. A small discrepancy between the feedback and overall self-assessment would indicate that the children are better at self-monitoring. In this case it is hypothesised they are applying the criteria for group skill usage consistently and observing incongruities (Johnson-Laird 1986). Stage 4 is designed to encourage the children to determine how they could have improved group skill usage in order increase the feedback value. This process is often performed automatically by adults in reflection-on-action (Schön 1987).

In Stage 4 the child is responsible for interpreting the feedback generated by the DSGS v1 model. The interpretation of the feedback will require *appropriation*. Appropriation is an interactive process in which the learner and expert engage in a joint activity. It consists of: 1) the learner making a move, 2) the expert interpreting the action using her framework and giving feedback, 3) learning occurs when the learner tries to understand in retrospect what his action meant to the expert and therefore updates his own framework of the activity. Newman gives an example of appropriation occurring by a novice computer user. The user – the learner – wishes to end a computer session but is still in a text editor, his move is to type "logout". The computer – the expert – presupposes the user wished to give an editor command because of the open application and responds with an error message. Consequently the user is made aware that the application open impacts the effectiveness of commands (Newman 1990, pp.84-85). Ideally the expert can explore the misunderstandings with the learner, but the learner can appropriate the feedback without this support (Newman 1990, p.88). The feedback generated by the DSGS v1 model is as independent of the learner's intention as in the example of the error message to the move "logout". This means it is possible for appropriation to occur at Stage 4, i.e., the child tries to create a framework that relates his actions to the mark given by the computer for group skill usage.

The claim that children can appropriate an expert model without being directly taught is supported by a simple example of appropriation observed in School X. After a design lesson one child voluntarily stayed to clean the classroom. This action was influenced by the state of the desks. Without explanation the teacher (the expert) awarded a "smiley", a reward for good

behaviour. This direct reward from the teacher can be seen as operant conditioning; however, it also influenced the behaviour of other learners. In the next design lesson more children remained to tidy up without explicit instruction from the teacher. These children appeared to have reflected upon how the feedback related to the action, which in turn influenced their understanding of the expert's model and hence their future behaviour.

Although appropriation can occur, knowledge of the expert's feedback alone may be insufficient to impact future behaviour. Newman *et al.* warn: "The appropriation process assumes that for any particular episode involving a novice [a learner] and an expert, the novice's psychological functions constitute an organized system that permits the novice to form some notion of what the episode is going to be about" (1989, p.64). In this training scheme this means the learners are required to understand the purpose of the feedback *and* be able to generate an understanding of the feedback that corresponds to the expert model.

Unlike a teacher the training scheme cannot assist in the construction of an "organised system" through monitoring the impact of the feedback to assist the child. It does not have the same access to the conversations, gestures, and knowledge of the participants. This absence of assistance can allow misconceptions to be generated and internalised by the children. From a Vygotskian perspective the computer implementation of the scheme has not assessed, however crudely, the current ability of the children. This means it is not working within their Zone of Proximal Development (Vygotsky 1978). The computer cannot act as a tutor or more able peer and organise the activity, leaving the children to construct an understanding alone introducing the possibility of error.

Instead, the training scheme works at two levels. First, as a tool it constrains the children's actions, the method of recording behaviour is restricted. At the second level, resources, the children's descriptions are influenced by the feedback given by the model. The feedback is in the form of prompts to think about group skill usage during the task and a request at the end to relate their actions to the numeric value generated (Sharples *et al.* 2002).

Requiring the children to work together enhances the possibility of appropriation in Stage 4. This means more exchanges are externalised as the group converses at the levels of actions *and* descriptions (Sharples *et al.* 2002). This matches Laurillard's (1993) minimum requirement for

learning to occur; the learner adapting his actions to his thoughts or descriptions based on the environment. In this situation the child must interpret information from two sources: the model's feedback and the understanding of fellow group members. The act of having to justify their actions and reasoning to their partner requires verbalisation, which helps the individuals to understand the practical task (Vygotsky 1978). Furthermore, it enables the knowledge structures of others to be integrated into the discussion with oneself *and* these discussions can result in the production of meaning (Smagorinsky 1998). These discussions are more effective if the justifications are elaborated rather than simple answers (Webb 1991, 1997). The model for this process is shown in Figure 4.3.

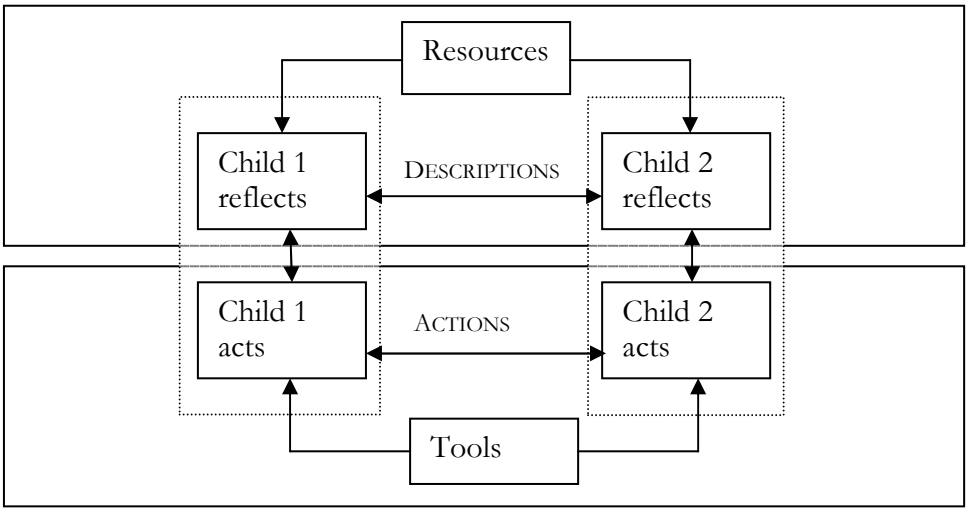


Figure 4.3: Appropriation process based on Sharples et al. (2002)

Giving responsibility to the children for appropriation does not mean the teacher is redundant with the introduction of the training scheme. The model can only provide prompts and lacks the historical and personality data that is needed to guide the interactions. It cannot ensure the interpretation the original expert model maps to the interpretation of it made by the child. By acting and reflecting in parallel with her pupils, a teacher can assist the children's understanding by tailoring the feedback (Laurillard 1993). She assists appropriation by creating an environment in which the children are exposed to her understanding (Newman *et al.* 1989).

4.1.5 Summary of the DSGS v1 model

The DSGS v1 model incorporates all three of the key attributes for working well as a group identified in the literature review (Section 2.3): reflection, responsibility and communication. This is summarised in Table 4.1. Its aim is to make the children evaluate their own group skill usage instead of providing a list of rules to be followed or giving the teacher responsibility for monitoring and evaluation. It uses techniques that have been shown to be effective in changing behaviour both for adults and children. The model is designed to make children aware of their interactions throughout the task *and* at the end to develop and support group skills.

SKILLS	METHOD OF INCORPORATION INTO THE DSGS V1 MODEL
Communication *Behaviour-based interactions*	• Prompts for group skills and criteria for assessment (Stage 1) • Reminders to discuss group skills during task (Stage 2)
Responsibility	• All individually being responsible for assessing performance during the task (Stage 2) • All individually being responsible for assessing performance at the end of the task (Stage 3)
Reflection	• Selection of appropriate behaviour and criteria for assessment (Stage 1) • Self-assessment during the task, overall and in appropriation stage after the DSGS v1 feedback is given (Stages 2, 3 and 4)

Table 4.1: Methods used to develop group skills

4.1.6 Requirements for using the DSGS v1 model

As shown in Figure 4.2 and discussed in the previous section, the teacher should monitor and support the groups during the GST. This is because the model cannot ensure the children are interpreting, or even, discussing criteria, group skill usage and the feedback. Her involvement is not the only factor though, that impacts the effectiveness of introducing the DSGS v1 model into the classroom. The concept of activity settings highlights that mediating artefacts are only one influence on learning (Engeström 1999c, Gallimore and Tharp 1990, Lipponen 2002). The DSGS v1 model acts as the "mediator of learning" by making the children's thinking about group skills explicit so they can be used in a new context (Adey and Shayer 1994, p.68). To be effective the model needs to be related to multiple domains, that is, the GST should be run concurrently with a variety of TGT, as in Talk Lessons or Jigsaw activities (Dawes *et al.* 2000, Aronson 2000).

Ideally, as well as in multiple contexts, the use of the DSGS v1 model requires a longitudinal component. Studies of teaching schemes to improve cognitive skills through raising awareness

have shown improvements are incremental over time (e.g., Dawes *et al.* 2000, Palincsar and Brown 1988, Adey and Shayer 1994). The DSGS v1 model recognises this: Stage 4, the feedback, influences the selection of rules, the definition of criteria and performance in future tasks.

The use of the model is also influenced by the expectation of the children. Study 1 indicated that the focus of the classroom is on domain skills. Despite the National Curriculum requirements (DfES 2000), this view is reinforced by the attitude of the teachers. This may influence the children's attitude to the tasks and explain the focus on the TGT. This focus may also be due to the metacognitive ability of the children. The GST requires the ability to reflect on the behaviour of oneself and the group. This is an abstract task and requires a certain vocabulary, which may not have been developed despite the discussions to construct it (Azmitia 1998, Sharples 1999).

This process of separating the two tasks may also cause problems in self-assessment. Children are not used to self-assessment and use criteria relating to task completion, participation and past and present teacher comments (Weeden *et al.* 1999, Study 1b). The GST is not related to the outcome of the TGT but the processes that occur *during* the activity. Children are required to separate these issues and focus on the group general learning skills. Similar higher order thinking skills are needed to interpret the feedback. It is not a summative score. The feedback highlights discrepancies in the application of criteria for group skill usage that need to be resolved. It proposes that the child should achieve this by appropriation, in which the children discuss and reflect on how the feedback relates to their actions. To ensure that they have sufficient vocabulary and awareness of actions, the previous GST stages must have been performed to the best of the children's ability.

4.2 Implementing the DSGS v1 model

As shown in Section 4.1 each stage of the DSGS v1 model has a theoretical basis which existing research has shown benefits group skill usage. This section describes how it was implemented on a computer. The computer used was of the specification found in the average primary classroom, and it need not be networked (BECTa and NAACE 1997). It discusses the advantages and disadvantages of using a computer as the medium and how the software corresponds to the model.

4.2.1 Advantages of using a computer as the medium

The use of computers as a mediating artefact for group activities is not new. Even without specialist software or hardware, research has found that computers encouraged talk (e.g., Light 1993, Wegerif 1996b, Hoyles *et al.* 1991b). Regardless of how groups are perceived, as individuals working together or as a joint unit, this social interaction can lead to cognitive development (Dillenbourg 1999). By making the computer the tool for both tasks the activity settings for the TGT and the GST can be overlapped. The TGT and GST are distinguished by their outcomes (Kuutti 1996). They share the same subject, community and mediating artefact and these, in the form of personalities, ability, group size etc., impact the outcome. The mediating artefact however, influences each of these factors. The computer as the mediating artefact in the GST gives responsibility to all group members in agreeing self-assessment and producing procedural rules for working together. The rules for the TGT are influenced by the activity, for example, whether the goal is to produce a written piece of work or find some information. This also influences the division of labour - how the task could be divided, whether a leader is required, etc. This is shown in Figure 4.4.

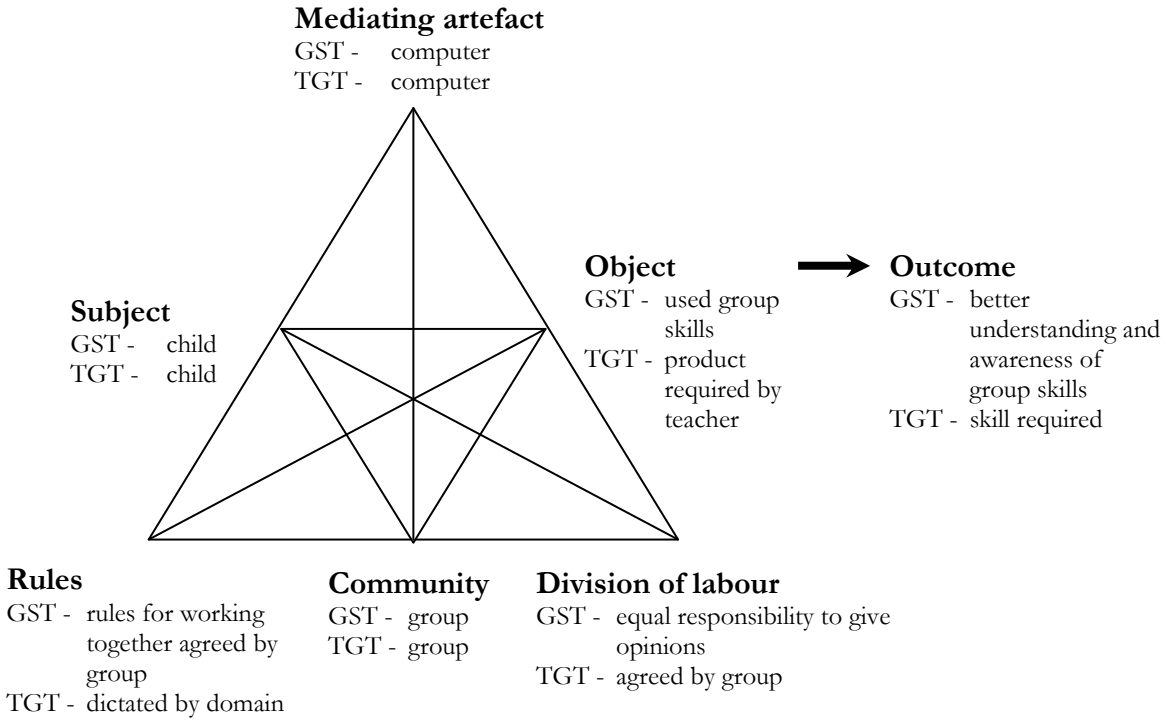


Figure 4.4: Activity setting for the GST and TGT, which run in parallel (based on Engeström 1999c)

In a classroom situation, "computers cannot have direct access to the collaborative activity itself" (Howe and Tolmie 1998, p.224). A computer program however, has indirect access. It can initiate and pace the task, monitor external aspects such as pauses and if being used usage of the dual-key control, and take account of student self-reported activities. A short pause between key presses, for example, indicates that a discussion has not occurred and this can be flagged to the participants. This suggests that even in face-to-face activities computer programs can influence the interactions within the group.

The DSGS v1 model to support the GST can be related to existing software structures. In Stage 1, a list of possible rules to be discussed can be implemented as a database. In Stage 2, prompts for discussion can be given. Even if these are independent of the TGT, they can result in exchanges leading to the generation of explanations, which improves understanding (e.g., White and Frederiksen 1998, Lester *et al.* 1997, Anderson *et al.* 2000). Prompts seed such conversations. Often pausing the activity generates more discussion than by written instructions or formally being told to talk by a teacher (Enyedy *et al.* 1997, Wegerif 1996b). Stages 2, 3 and 4 in the DSGS v1 model require information to be logged, prompts given on the basis of these values, and then the information re-represented as feedback. These practices are common in "drill and skill" packages as well as more complex tutoring programs.

4.2.2 Disadvantages of using a computer as the medium

Implementing the DSGS v1 model using a computer as the medium raises further issues. The domain of the TGT is limited to those in which a computer is used. Furthermore, the preconception of computers held by teachers and pupils will influence the effectiveness in raising and developing awareness of group skills. The National Curriculum Key Stage 2 activities for Information and Communication Technology (ICT) focus on practical skills, such as using word processors, spreadsheets and databases (DfES 2001). If teachers see ICT lessons purely as a curriculum subject, then it may be harder for them to see such activities as a method of developing group skills. Since children reflect the attitude of their teachers, they may also focus on the TGT, rather than using the computer as a mediating artefact for the support and development of group skills (Study 1, Gallimore and Tharp 1990).

A further difficulty is the expectation of information entered into a computer and how feedback is treated. Study 1d showed the most common use of the computer, apart from Internet

searching or word processing, was "drill and skill" packages for mathematics. In the latter, the children are required to give the correct answer, and at the end, they receive summative feedback. Implementing the GST on a computer requires the children to perform a novel activity – recording skill usage on the computer. Although the recording is subjective, the format of providing feedback at the end of the activity is similar to existing activities relating to the correctness of the input. Then there is the issue of how this feedback is used. When the activity was to interrogate the Internet, the children viewed the task as calling up the web page, not reading and understanding the information it contained (Study 1d). Like a web page, the feedback produced could be seen as the product, rather than a piece of information to be internalised.

4.2.3 Implementation of the model

As discussed above there are advantages and disadvantages of using a computer as the mediating artefact for the model. This section considers one specific computer implementation of the model designed to investigate whether the children can perform the two activities simultaneously. This is to determine whether the GST has a detrimental impact on the TGT performance and if the children can self-monitor in the GST. It also investigates whether this information is sufficient to generate a feedback value that accurately reflects group skill usage. It begins by describing the TGT and its implementation. Section 4.3 describes the study in which this software was used. The remaining issues are addressed in later chapters. Chapter 6 investigates the impact of using a computer as the medium. Chapter 8 describes the impact observed with repeated usage.

The software was designed for pairs due to the following theoretical and practical considerations:

- Dual-key control could be implemented forcing communication between the children and engendering a degree of responsibility as both are accountable for any decisions (Stewart *et al.* 1999, Light and Glachan 1985)
- Children in Study 1a recommended small groups so they can physically participate in the activity
- Teachers in Study 1c see computers as a practical activity and are used to children working in pairs at a computer
- There is more opportunity for talk (Puntambekar 1995)

- Working in pairs at a computer results in a significantly higher level of task engagement (Inkpen *et al.* 1999) and is motivating (Stewart *et al.* 1999, Scott *et al.* 2002)
- The number of available computers and classroom layout meant children had to share a computer

A problem-solving task was chosen as the TGT (Study 1c, Galton and Williamson 1994, p.43). To extend the work of Light and Glachan (1985) the Tower of Hanoi puzzle was used. It has a clearly defined domain, is easy to represent, requires minimal writing by the children, and the number of moves can measure task performance. Since the software incorporates the DSGS v1 model and the task, it is not domain independent. The software can record task performance and the self-monitoring.

The TGT was presented in the form of an adventure game called The Escape from the Forbidden Forest. It was written in Macromedia Director 7. The story, screen design, interaction techniques, heuristic to calculate the time spent using group skills from the self-monitoring stage, and questionnaires were developed with five teachers and eight children of differing abilities from Schools X and Y (Druin 2002).

The TGT begins by explaining how the pair of children and their two friends came to be stranded up a tree in the Forbidden Forest. To be rescued the children must move all four characters, represented by different size blocks, from tree 1 to the faded tree 4. For the characters to be safe from various ghosts, sorcerers and man-eating monsters, only one of them can be moved at a time. The puzzle was implemented as a Single Display Groupware (SDG) system. Dual-key control was integrated as one child selects the character using the mouse, the other types in the tree using the keyboard. Once a tree is selected, the chosen character automatically climbs up as high as possible. If a character MUST climb over another to reach the highest branch that can bear his weight then the move is disallowed and a life is lost. This is because the combined mass of the two characters would cause the branch to snap and the characters to fall and be attacked by the forest inhabitants. If all five lives are lost, the puzzle is restarted. (The four trees equate to the rods in the Tower of Hanoi problem, and the four characters to the rings.) Help about the task and screen display can be found by clicking on the question mark button on the bottom right of the screen, and the number of lives and moves taken are also shown.

After the first successful solving of the puzzle, the children can choose to continue practising or make their final attempt. This involves solving the puzzle alone, with the start and end trees randomly selected so the problem is not identical to the practice task. By telling the group both must be able to complete the task in the minimum number of moves responsibility is incorporated into the practice stage. The children communicate to make a move, but must also reflect and communicate about the TGT to ensure their own understanding and that of their partner. This design was meant to encourage collective or collaborative working, that is, to be a joint activity–joint outcome task.

The remainder of Section 4.2 describes how the puzzle, The Escape from the Forbidden Forest, incorporates the four stages of the DSGS v1 model.

4.2.4 Stage 1: Raising awareness

The children use their knowledge, teacher support and a database within the software to select or invent three rules for working well together. They can enter their own using the sample rules as prompts, or select an existing rule by double clicking. At this stage, they are prompted to discuss the assessment criteria. They can say the "Agree before moving" rule, for example, has been followed if each gives the other the chance to say no or ask why a move should be made and not when one simply tells the other what to do.

The database was compiled by the teachers and pupils who participated in the design phase. They selected rules from existing research (Study 1, Mercer *et al.* 1999, Wegerif 2001, Kagan 1992). The sample rules chosen were:

- Listen to each other
- Take turns
- Make sure you both understand
- Agree before moving
- Do not fight
- Help each other
- Do not cheat

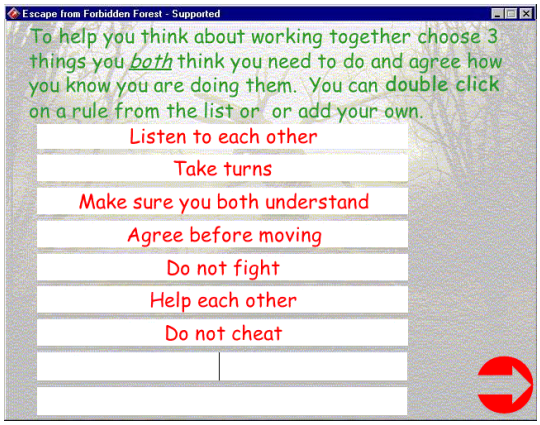


Figure 4.5: Stage 1 screenshot from The Escape from the Forbidden Forest

The database is small due to the screen size and amount of information a pupil is willing to read. The interface is shown in Figure 4.5. Some rules were influenced by practical considerations. "Do not cheat", for example, was included as the computers were adjacent so copying was a temptation.

4.2.5 Stage 2: TGT performance and self-monitoring

The chosen rules are displayed on the screen throughout the task, as shown in the top left corner of Figure 4.6. They act as reminder for appropriate behaviour (Dawes *et al.* 2000) and are used in the monitoring activity. At the start of Stage 2 the children are asked to monitor their behaviour. If they *both* agree a rule if it had been followed while making a puzzle move they are asked to click on that rule. If they disagree, or both agree a rule has not been followed, they are asked to discuss why and plan how it should be followed in future. The child with the mouse has control over what information is entered. To indicate that a rule has been selected there is a beep and the rule colours are inverted. Rules can only be selected once between puzzle moves. If after five TGT moves (moving characters between trees) no rules have been selected a simplistic procedural prompt is given. This takes the form of a message box reminding children to monitor their behaviour and click on the rules when appropriate. The software stores the number of prompts and the rules selected in the "Record of self-monitoring".

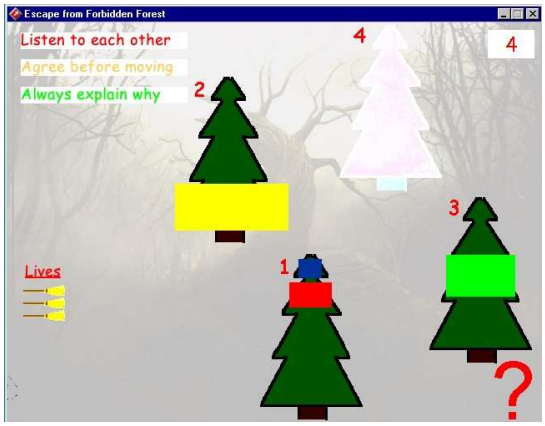


Figure 4.6: Stage 2 screenshot in The Escape from the Forbidden Forest

4.2.6 Stage 3: Overall evaluation

On completion of the TGT, the children are asked to discuss how well they believed they worked together, including providing examples and explaining their reasoning. They then quantify and input this result using a scale of 0 – "not worked together at all", to 100 – "worked perfectly" as shown in Figure 4.7. This is stored in the "Record of self-monitoring" by the software.

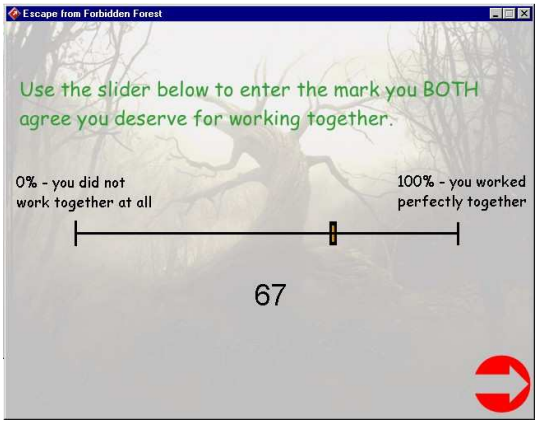


Figure 4.7: Stage 3 screenshot in The Escape from the Forbidden Forest

4.2.7 Stage 4: Feedback

To generate a heuristic that created feedback similar to that of a teacher, the teachers assessed videos of the children involved in the design stage working together at the "escape". They were given a three-minute clip of each pair and asked to: i) individually assess the time spent using group skills on the same scale as the children in Stage 3, and ii) after as a group to discuss their opinions. The scores were consistent within the schools, differing by, at most, 5% in all cases. The discrepancy between the schools was larger in one case; those in School Y gave one pair a mean mark that was 27.5% lower than School X teachers. The discussion implied this was due to the volume of noise these children produced. Although on task they were excited because they were about to complete the TGT. All the teachers used similar criteria. Body language was seen as the most important factor for judging the amount of group work. Children both looking in the same direction and taking turns to speak indicated a higher level of group working. Gazing in different directions, or talking simultaneously indicated a low level. The second factor is the children's conversation. Teachers listened for phrases used in explanations, and in building consensus, such as "because" and "do you think we should…". Finally, requests for help without attempting the problem first were seen to indicate a lower level of group working. In contrast to Study 1c, the teachers questioned would not use task performance as an indication of group skill usage, nor would they incorporate the children's own opinion – despite Ager (1998) recommending the inclusion of both.

Given the scores and comments from teachers reviewing the videos, a heuristic was constructed that generated a similar value using the information available (Appendix H1 gives the heuristic and an example of the responses generated). This formula used the:

- Frequency the children recorded following a rule using their agreed criteria
- Number of reminders given
- The similarity between the individual solutions (this information is included since if the children worked effectively together their solutions would be similar (Light 1993))

Gallimore and Tharp state: "Simply providing performance information is not feedback, there will be no performance assistance unless the information provided is compared to some standard" (1990, p180). Representing the feedback as a number ensured that it was easily understood; children recognise that a higher number relates to a higher standard. Depending on the difference between this "score" and the opinion entered in Stage 3, the software prompts the

children to discuss and to determine an explanation for why the mark is higher or lower, and if all marks are high then it asks why, and if not what could be done to improve the scores in the future. An example is shown in Figure 4.8.

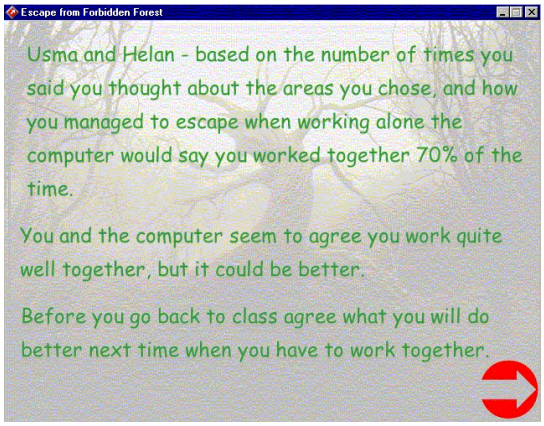


Figure 4.8: Stage 4 screenshot from The Escape from the Forbidden Forest

4.3 Study 2: Testing the DSGS v1 model and software

As stated, the first implementation and study using the DSGS v1 model did not seek to address all the issues raised in the design and development phase. The goal was to test whether:

- Performing these two group tasks was detrimental to task performance
- Self-monitoring was possible when done concurrently with a group task
- Simple self-monitoring provided enough information for a "score" of time spent using group skills to be generated

The remainder of this chapter describes the study to address these issues. The first point was examined by comparing the task performance between children using the model and those working alone or in pairs without this support. For the second issue data from the self-monitoring, a questionnaire, and the teacher's opinion were used to establish the children's perception of group work and ability to self-monitor. Finally, to indicate if a simple self-monitoring system is sufficient, the "score" of group skill usage generated by the software was correlated with the teacher's opinion.

4.3.1 Participants and setting

In total, 146 children aged 9 and 10 participated. They had not taken part in Study 1. Two classes came from School X, a Church of England School in an affluent area of Birmingham, 12% have English as a second language and 22% are from ethnic minorities. Three classes came from School Y, a city primary school in Birmingham. In this school, half the children come from ethnic minority backgrounds and have English as a second language. In the following year, the children who participated from School X performed above the National Average in the Standard

Assessment Tests, by 17% in mathematics and 15% in English. The participants from School Y performed below the National Average in the same tests taken at 11 in mathematics and English by 10% and 25% respectively. Both schools performed above the National Average in science, by 11% and 2% respectively. (See Appendix F for a comparison of the formal assessments for the schools.) All the children were familiar with computers since ICT lessons were compulsory from Year 1.

4.3.2 Method

The TGT was The Escape from the Forbidden Forest, based on the Tower of Hanoi puzzle and described in Section 4.2.3. The study took place in each class's ICT lesson. These were held in the school's computer room and lasted approximately 45 minutes. The lessons in both schools were held on the same day. Each room contained approximately 17 personal computers. As far as possible, the same number of machines was used for each condition.

The children were randomly divided into three conditions, each playing one variation of the task. Two of the conditions required the children to work in pairs, in the third the children worked individually throughout. Both conditions for pairs had dual-key control, one child used the mouse to select a character (the ring), and the other chose the numerically labelled tree (the rod) through the keyboard. In the third condition, the child had control of the mouse and keyboard. For a character to move both actions were required. In the *supported pairs* condition the puzzle incorporated the DSGS v1 model so the GST was explicit. The *non-supported pairs* condition is equivalent to the Light and Glachan (1985) groups - the children are told to work together, but are not given any support.

In the conditions for pairs, the children were allowed to choose their partner. According to the questionnaire completed at the end most worked with their best friend or a friend: 78% in non-supported pairs, and 80% for supported pairs. The rest worked with someone described in the questionnaire that they "got on OK with".

The only instructions given verbally to the children were how to access the software. Once the application was open an identical written introduction and set of rules to move characters was given to all three conditions. Before starting the TGT the children working in the supported pairs condition chose three "rules" and were asked to discuss assessment criteria before starting

the puzzle, as shown in Figure 4.5. At this point children working in non-supported pairs were asked by the application to work together. All conditions were then given a practice session. During the "escape" the supported pairs were expected to monitor their behaviour by clicking on the displayed rules. Failure to click on any of the rules generated a reminder after five TGT moves. After the first successful task completion, the children could choose to continue practising or make their final attempt.

In both conditions for pairs, after the children had individually completed the task they were asked to assess together how well they worked together overall using a scale of 0 – "you did not work together at all", to 100 – "you worked perfectly together". After their final attempt, children working independently were asked to assess how hard they found the problem, from 0 – "impossible", to 100 – "very easy". Each variation generated an opinion on this question using the same scale. In the case of the supported pairs the heuristic generated a score based on the: i) frequency the children recorded following a rule using their agreed criteria, ii) number of reminders given, and iii) similarity between the individual solutions. In the non-supported pairs the opinion is based only on the similarity of the individual solutions. In the individual variation, the opinion is based on the number of moves (Appendix H describes all the heuristics in full). The software asks those in pairs to discuss any discrepancies between their own opinion and the one generated and to try to give possible causes. Those working independently were asked to provide an explanation of why their opinion differed from that of the computer.

During the session the class teachers recorded on paper their own evaluation of how well the pairs worked together on the same scale of 0 to 100. As described in Section 4.2.7 the criteria were similar across both schools and the marking within schools was consistent.

At the end, the children anonymously completed a questionnaire according to the variation used. (Copies of the questionnaires and an analysis of the findings can be found in Appendix I.) All the children gave an opinion of the puzzle and how well they normally worked in groups at computer tasks. In addition, pairs recorded how they felt they worked together in this task compared to previous computer lessons. After the puzzle, some of the children who had worked in supported pairs were asked what they thought about the self-monitoring process.

Performance data was gathered from 118 children: 23 supported pairs, 23 non-supported pairs, and 26 individuals. 135 children completed the questionnaires: 51 from children in supported pairs, 60 in non-supported pairs, and 24 individuals. The discrepancies are due to some children not returning their questionnaires and one class having technical problems with the computers.

The results must be treated with a degree of caution. The questionnaire results show of those supposedly working independently, approximately 60% helped those around them when asked, and 10% offered advice without being requested. Approximately 35% of children in non-supported pairs and 40% of those working in supported pairs said they assisted their partner instead of working individually in the final attempt. Despite the opportunity for discussion during the break, there is no significant difference in task performance in later classes.

4.4 Task performance

This section discusses whether using software incorporating the DSGS v1 model is detrimental to the understanding of the task.

4.4.1 Results

The results show that despite help offered by others, children who worked in supported pairs completed the task in fewer moves. This is the case for the first attempt, where they worked together, and for the final attempt, where they worked independently. In turn, non-supported pairs used fewer moves in both cases than those who worked independently throughout. This is shown in Figure 4.9. There is a significant difference in TGT performance in the final attempts, with those who were supported requiring least moves, followed by those who were non-supported, and finally those worked alone throughout ($p>0.01$ using Kruskall Wallis).

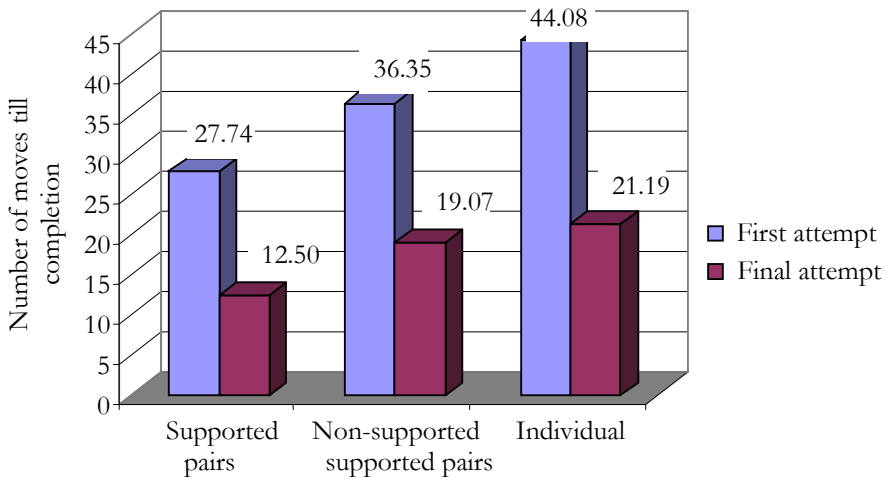


Figure 4.9: Average number of moves to "escape"

4.4.2 Discussion

Study 2 indicates that supported pairs have a better task performance measured by the number of moves compared to children working in non-supported pairs, or children working alone. This is despite 70% of children saying they helped those who were supposed to be working individually. These results strengthen research showing collaboration improves task performance (Farivar and Webb 1994, Ashman and Gillies 1997, Mercer *et al.* 1999).

It also extends the work of Light and Glachan (1985). They found an individual child needed fewer moves to complete the Tower of Hanoi puzzle if they had practiced in pairs using a dual-key control instead of being requested to work together as a pair. In this study, performance is further improved using the DSGS v1 model *and* dual-key control. This suggests using the DSGS v1 model improves task performance more than dual-key control alone. These results must be treated cautiously; testing for individual understanding immediately after the practice may not reflect the actual cognitive skill gained (Howe 1991).

4.5 The children's perception of collaboration

In this section the ability of the children to self-monitor is examined.

4.5.1 Results

In the questionnaire the children were asked how well they normally worked together when using a computer and in this particular TGT. The responses are shown in Table 4.2.

	QUESTIONS			
	1. How much working together do you normally do on a computer?		*2. How much working together did you do when trying to escape together?*	
SELF-ASSESSMENT ANSWERS (5-POINT SCALE):	**Supported pairs**	**Non-supported pairs**	**Supported pairs**	**Non-supported pairs**
Always agree what to do first	58.8%	30.0%	41.2%	31.7%
Usually agree what to do first	15.7%	28.3%	41.2%	26.7%
You take turns and help sometimes	11.8%	33.3%	11.8%	33.3%
You take turns and help if asked	11.8%	6.7%	5.8%	8.3%
You take turns but do not talk about the problem	1.9%	1.7%	0.0%	0.0%

Table 4.2: Questionnaire results showing the children's perception of the amount of collaboration.

In both conditions for pairs a significant number always or usually agreed they worked well together both normally when using a computer and for this specific TGT ($p>0.001$ using Chi square). From question 1 significantly more children working in the supported pairs condition believed they always agreed what to do when using a computer compared to children in the non-supported case ($p>0.01$ using Chi square).

A comparison of the two conditions for question 2 shows supported children stated they worked together more when "escaping" than those who were *not* supported ($p>0.1$ using Chi square). Over 82% who worked in supported pairs judged themselves as always or usually agreeing what to do first, compared to around 58% in non-supported pairs (see Table 4.2). However, there is a decrease in the number of children who stated they always agreed what to do first in the supported pairs condition. This fell from 58.82% in normal computer sessions (question 1) to 41.18% when solving this puzzle (question 2).

An analysis of the self-monitoring of supported pairs showed the agreement and turn taking recorded in the questionnaire was not reflected in the records created during the task. The average volume of self-monitoring by the supported pairs condition in the ICT lesson was approximately three times less than that recorded by the children in the design stage. In addition, ten out of the 23 supported pairs had more reminders to consider the rules than the number of times they recorded following them. Discussions with supported pairs after the puzzle indicate low levels of self-monitoring were due to absorption in the TGT.

There is no relationship between the marks given by the teacher and the overall self-assessment marks the children awarded themselves in either the supported or non-supported conditions. Teachers assessed non-supported pairs as displaying significantly more group skills than supported pairs, giving an average of 71.3% and 58.0% respectively ($U=166$, $p>0.05$ using Mann Whitney). A breakdown of the children's own overall self-assessments are found in Table 4.3. In contrast to the answers to question 2, more children in the non-supported pairs believed they used group skills over 90% of the time. Additionally, more children from the supported pairs believed they used group skills less, and more in the supported pairs gave themselves a non-zero value below 50% as an overall assessment. The two pairs who assessed themselves as using group skills under half the time in the non-supported condition said they never worked together.

	TEACHER'S MEAN (%)	CHILDREN'S MEAN (%)	NUMBER OF PAIRS WHO GAVE THEMSELVES LESS THAN 50%		NUMBER OF PAIRS WHO GAVE THEMSELVES MORE THAN 90%	
			OUT OF 23	%	OUT OF 23	%
Supported pairs *Standard deviation*	58.0 *15.37*	76.7 *30.27*	3	13.0	9	39.1
Non-supported pairs *Standard deviation*	71.3 *25.53*	78.3 *30.73*	2	8.7	12	52.2

Table 4.3: Comparison of group assessment scores for working together on task completion

4.5.2 Discussion

The questionnaire findings support the Study 1a results. The questionnaire and overall assessment show that the children participating have a high opinion about the amount of group skills used. The insignificant difference in the amount of working together done in normal computer lessons and in the puzzle The Escape from the Forbidden Forest recorded by the non-supported pairs in Table 4.2 suggests children do not believe that dual-key control has an impact on behaviour. This is despite the finding that dual-key control encourages communication (Stewart *et al.* 1999) and it results in better task performance (Light and Glachan 1985). Children working in supported pairs were significantly more likely to state they always or usually agreed compared to non-supported pairs (question 1 from Table 4.2). There was still, however, a significant increase in the number that felt they worked better together when solving The Escape from the Forbidden Forest (question 2 from Table 4.2). This indicates that it is the process of being asked to *choose and reflect on rules for working together* that causes children to feel they collaborate more, rather than the dual-key control.

The low marks for group skill usage given by teachers to supported pairs may be related to the amount of talk observed. Teachers in School Y gave a low mark to children who were excited and talking about the TGT. The DSGS v1 model encourages talk. The children are required to discuss criteria for assessing group skills, give examples, plan future behaviour, as well as negotiate which character to move in the puzzle. If talk is seen as disruptive rather than constructive, despite being necessary for developing an understanding, it may impact the teacher's assessment of the amount of time spent using group skills.

The process of reflection appeared to impact on the children's opinion of the level of group work displayed, despite the lack of input from the teachers during the TGT. From Table 4.2 the

number of children who believed they worked perfectly while performing The Escape from the Forbidden Forest was 17% less than in standard computer lessons. This is supported by the self-monitoring in Stage 3. Approximately the same proportion believed they worked together all the time in answer to question 2. There are discrepancies, as due to the technical difficulties the same children did not provide the data in Table 4.2 and Table 4.3. However, the fact four children in non-supported pairs believed they never worked together contrasts with no children stating this in the questionnaire. Although not significantly different values, children in supported pairs appear to be more consistent than non-supported pairs. Children who did not monitor behaviour throughout the tasks (non-supported pairs) gave themselves more higher assessments and fewer lower ones (Table 4.3). These beliefs do not match the proportions from the questionnaire (Table 4.2).

The decrease in the number who always agreed what to do first in the supported pairs between questions 1 and 2 could suggest the children believed they did not always use groups skills during the TGT. This could partly explain the low monitoring levels in Stage 2, as children may have agreed they were not following any of their chosen rules rather than the given explanation of being involved in the TGT. Enyedy *et al.* (1997) offer an alternative explanation. They found that prompts caused discussion amongst the children leading to a clarification in understanding, although the children ignored the request to record their views. Another explanation for the lack of monitoring is the context. Unlike a laboratory-style setting, as in the design phase where the children were aware of being observed and knew the goal was to raise awareness of group skills, the classroom context in Schools X and Y is geared towards the TGT. This is corroborated by the interviews with the children, who stated the prompts were ignored because of involvement with the puzzle, although it is possible the children failed to read the reminders on the prompts or the child with the mouse dominated the task and overruled their partner. Regardless of reason, the low amount of monitoring compared to the design phase meant it was impossible to determine whether this form of continual self-assessment in a classroom context could raise incongruities and reduce the overestimation in group skill usage.

Assuming the low level of self-monitoring is due to the classroom context and not a lack of metacognitive ability preventing the children switching between the GST and TGT, these findings suggest refinements to the DSGS v1 model. The current approach of recording whether a rule was followed does not reflect the self-monitoring that occurs. If the children do not

record using a group skill it may be because they are absorbed in the TGT and forgot this was required, or did not read the prompts, or considered it but agreed it was not used, or felt it was used insufficiently to be recorded. To overcome the issue of forgetting or not reading the prompts the self-assessment could be made compulsory. The later issues would be overcome if a scale were provided so the children could record the proportion of time a rule was followed (Kagan 1992). Such an approach would pause the activity and if prompts were provided could encourage communication on group skills (Enyedy *et al.* 1997, Wegerif 1996b). Finally, to prevent domination by one child both children should be given responsibility for entering their own opinions.

4.6 Generating feedback

This section looks at the last question, whether recording each time the group believed a chosen group skill was followed is sufficient to generate a "score" for working together.

4.6.1 Results

The opinion of the teacher was plotted against the result generated by the heuristic in the software incorporating the DSGS v1 model, as shown in Figure 4.10a. There was no significant correlation between the two (R = -0.312). Figure 4.10b shows the correlation between the score generated by the system based on task performance used by non-supported pairs and the teacher's opinion. Again there is no significant correlation (R=0.128).

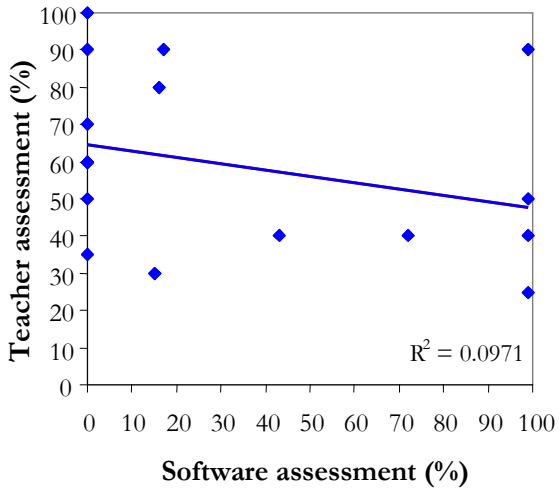
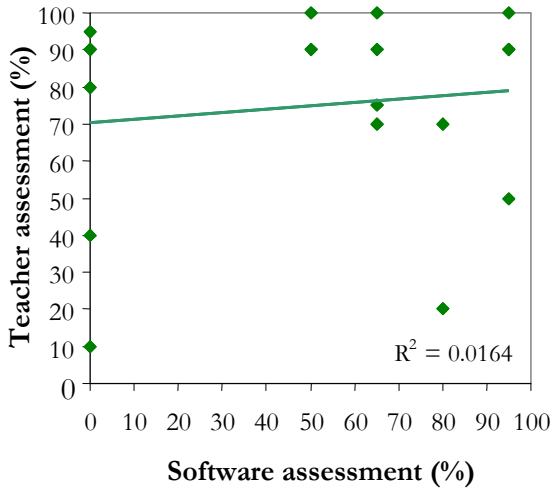


Figure 4.10a: Correlation between the DSGS v1 software heuristics and the teacher's assessment of group work in supported pairs



b: Correlation between the score generated on task performance and the teacher's assessment of group work for non-supported pairs

Zero marks were given to non-supported pairs if there was a large discrepancy between the number of moves taken when working individually. In supported pairs, although this factor plays a role, the high number of null scores for the assessment of supported pairs was due to: i) infrequency in recording if a chosen criterion was followed, and ii) a high number of reminders (see Appendix H1 for the heuristic).

4.6.2 Discussion

The initial viewing of videos of children working together before the study, showed the teachers believed task performance alone is an inadequate measure of the amount of group work demonstrated. They agree with the opinions of Crook (1998) and Mercer (1994). It is individual personality, experience and the personal style of working that affect task performance as much, if not more, than the knowledge gained through working together. This is supported by Figure 4.10b. This shows there is no correlation between the teacher's opinion and an assessment based on task performance and the similarity in the solutions. This is also reflected in teachers assessing non-supported pairs as displaying more group skills, despite taking more moves to solve the task as shown in Figure 4.9. This supports the idea that the software incorporating the DSGS v1 model can be domain independent. The task performance is not indicative of the amount of group work displayed. Separating the activities would also be beneficial if the group skills are to be easily transferred. As Salomon (1995) warns children may perceive the activity, in this case self-monitoring as integral to the task of escaping from the Forbidden Forest, rather than as an independent skill.

The lack of correlation in Figure 4.10a shows the current heuristic cannot generate an opinion similar to the teacher's. Task performance and the children's voluntary recording of group skill usage in a classroom context provide insufficient data. This was not an issue in the laboratory-style setting used in the design phase. In this environment, however, the children's behaviour may have been affected by the obvious monitoring and video camera (Biott and Easen 1994b) and the researcher's presence (Pilkington and Jones 1996). In the classroom setting there were none of these influences. This suggests that the procedural prompts should be expanded to take the role of the adult, prompting for discussion on why opinions are held as well as reminders to the children to reflect on behaviour.

4.7 Summary

In this chapter a model to develop and support group skills (DSGS v1) was presented. It proposes that a group should perform two activities concurrently. The first is the group activity that would be given by the teacher (TGT). The second explicitly encourages the children to monitor their interactions so that their everyday understanding of group skills can be abstracted and applied in multiple environments. This is the group skills task (GST). The methods employed are based on existing training schemes. The DSGS v1 model focuses on raising awareness of communication, reflection and responsibility.

The advantages and disadvantages of a computer as a medium for the implementation were then discussed. Computers encourage discussion and are motivating. The four stages within the model: awareness raising, prompting for self-monitoring during the TGT performance, prompting for evaluation of overall group skill usage, and providing feedback, can easily be implemented. As when using any training scheme, the outcome is dependent on the context. Using a computer as a mediating artefact for group skills must overcome existing expectations. Children and teachers may view ICT as a curriculum requirement and a set of skills to be obtained. Feedback generated by a computer is often a summative score, based on answering the computer prompted questions. Information returned could be seen as a product in itself rather than being appropriated through discussion.

Although these issues are important, the chapter concludes by discussing a study using one implementation of the DSGS v1 model. Study 2 suggests that incorporating the TGT and DSGS v1 model into one software package so the TGT and GST are run simultaneously is beneficial to the TGT understanding. Children reflecting on criteria for collaborating at the start, and being prompted for self-monitoring during the task - even if such monitoring does *not* occur, complete the TGT in the least number of moves. It also supports existing research findings that children working in groups achieve better task performance than individuals.

Having to consider group skills before a task commences and monitor interactions during a TGT results in the children believing they work better together. This contrasts with the views of the teachers who believe less group skills are shown in this situation. One suggestion for this incongruity is the noise generated by the supported pairs. The DSGS v1 model encourages talk, which can be seen as disruptive. Although the average overall assessment of group skill usage is

high, many children have difficulty in self-monitoring during the task. Used within a classroom environment a simple monitoring system provided insufficient data to generate a "score" for working together. However, such a score is not dependent on the TGT performance.

A major drawback to this study was the passive role of the teachers, which appeared to lead to the lack of self-monitoring and discussion in the classroom setting. The model requires the teachers to remind and assist the children in monitoring their behaviour and developing definitions of group skills. Without teacher participation children may fail to develop accepted definitions of group skills, as they cannot monitor and arbitrate their discussions. This role of prompting discussion could be partly taken by a computer; the procedural prompts could be expanded to draw attention to how their opinions are formed *and* that they should be reflecting on behaviour.

The study also indicates that the model is not domain specific, as task performance is not reflected in group skill usage. It also shows that the DSGS v1 model needs refining. Highlighting appropriate group skills in Stage 1 appears to benefit the TGT. This, and the requests to self-monitor during the task in Stage 2, suggest the overall reflection in Stage 3 is taken more seriously. The problem is the failure to record the monitoring of group skill usage in Stage 2. This could be overcome by:

- Making self-assessment compulsory
- Providing levels representing the proportion of time spent using a group skill
- Ensuring both children enter an opinion - in a computer implementation for pairs this could be achieved through dual-key control

Chapter 5 discusses how the DSGS v1 model was refined and the impact this has on the TGT and GST performance.

Chapter 5

THE DSGS V2 MODEL

In Study 2 the children appeared to focus on the teacher-given task (TGT) rather than on the group skills task (GST). The Developing and Supporting Group Skills (DSGS) v1 model was amended to ensure the GST was performed by making the assessments compulsory, using a scale of measurement based on the proportion of time a group skill was used, and requiring all the group members to enter an opinion. These refinements incorporate reflection and responsibility but not necessarily communication. Communication is important as before internalisation can occur the knowledge of group skills occurs on the interpsychological level (Vygotsky 1978). This can be achieved through discussion. To ensure that discussion occurs *and* the children develop a shared vocabulary, the DSGS v2 model uses the self-monitoring to provide prompts to focus the exchange. If the children believe, for example, a group skill was used for different proportions of time it would require them to explain why. The prompts also scaffold behaviour within the TGT. An example is the explicit request to plan and implement a method to improve usage of a group skill if the children record in the GST self-assessments it is being rarely used.

In the first section, Section 5.1, the differences between DSGS v1 and DSGS v2 are listed and the theoretical basis for the changes are discussed. One possible software application - Aliens, is described in Section 5.2 to illustrate how the model could be implemented. The software builds on the assumptions of the suitability of computers as a medium identified in Sections 4.2.1 and 4.2.2. It was developed with children acting as informants, that is, users and testers (Druin 2002). This ensured the software was motivating, the format was clear, and that the concept of equal responsibility through the dual-key control could be understood.

The remainder of the chapter describes two studies using the revised model, DSGS v2. Study 3a addresses whether the incorporation of compulsory individual self-assessments impacts upon the TGT performance. If the training scheme to raise awareness of group skills is detrimental to the acquisition of the domain knowledge, then it will not be used within the classroom (Section 2.5 and Study 1c). The impact of the revisions was examined by incorporating the DSGS v2 model into the same puzzle used in Study 2. Here the number of moves taken to solve this TGT was compared to that needed when using the DSGS v1 model.

Study 3b investigates the reliability of the self-assessments. Studies 1 and 2 showed that the children overestimated the *overall* group skill usage compared to the adults observing them. This belief that they worked well together was also found in assessments *during* the activity. In Study 3a, the majority of children believed they used the group skill under consideration all the time. The feedback given to the group on the proportion of time they spent using group skills is based on these self-assessments. For the feedback to be accurate, the children must apply the criteria for assessing group skill usage consistently during, and at the end, of the TGT. To determine if the overestimation and inconsistency is due to the children using the vocabulary differently to more experienced members of their community, or failing to relate actions to their understanding of group skills, the children were asked to assess three scenarios of pupils working together.

5.1 Theoretical basis for model refinement

Three of the four stages were revised. The first stage in DSGS v2 is identical to that in the DSGS v1 model. In Stages 2 and 3, dual-key control was introduced allowing individual assessment of group skill usage throughout. A form of procedural prompting (Palincsar and Brown 1988, King *et al.* 1998 etc.) was introduced to ensure that communication occurred between the children. Discussion prompts were generated from generic question stems based on their self-assessments; depending on their opinions the children were required to justify, plan or give examples of group skill usage. The prompts also included a type of scripted co-operation (Burton *et al.* 1997, Goodman *et al.* 1998, etc.). The prompts randomly assigned roles to the children to ensure turn taking and that each child could generate explanations and evaluate the reasons given by others. The introduction of prompts introduces two principles identified by Tharp (1993) for improving cognitive development. The first is *questioning*. This focuses discussion allowing a shared definition of group skills and criteria for assessment to be agreed. The second is as a form of *modelling*. The prompts illustrate the type of exchanges that occur in an effective group. The actions of planning, justifying or giving examples in the GST are equally applicable to the TGT domain and could be transferred to discussions in that activity. This extends the idea in the DSGS v1 model that the GST and the TGT influence each other: the assessment process influences behaviour and the behaviour is reflected in the assessment.

The refinements are highlighted in Figure 5.1.

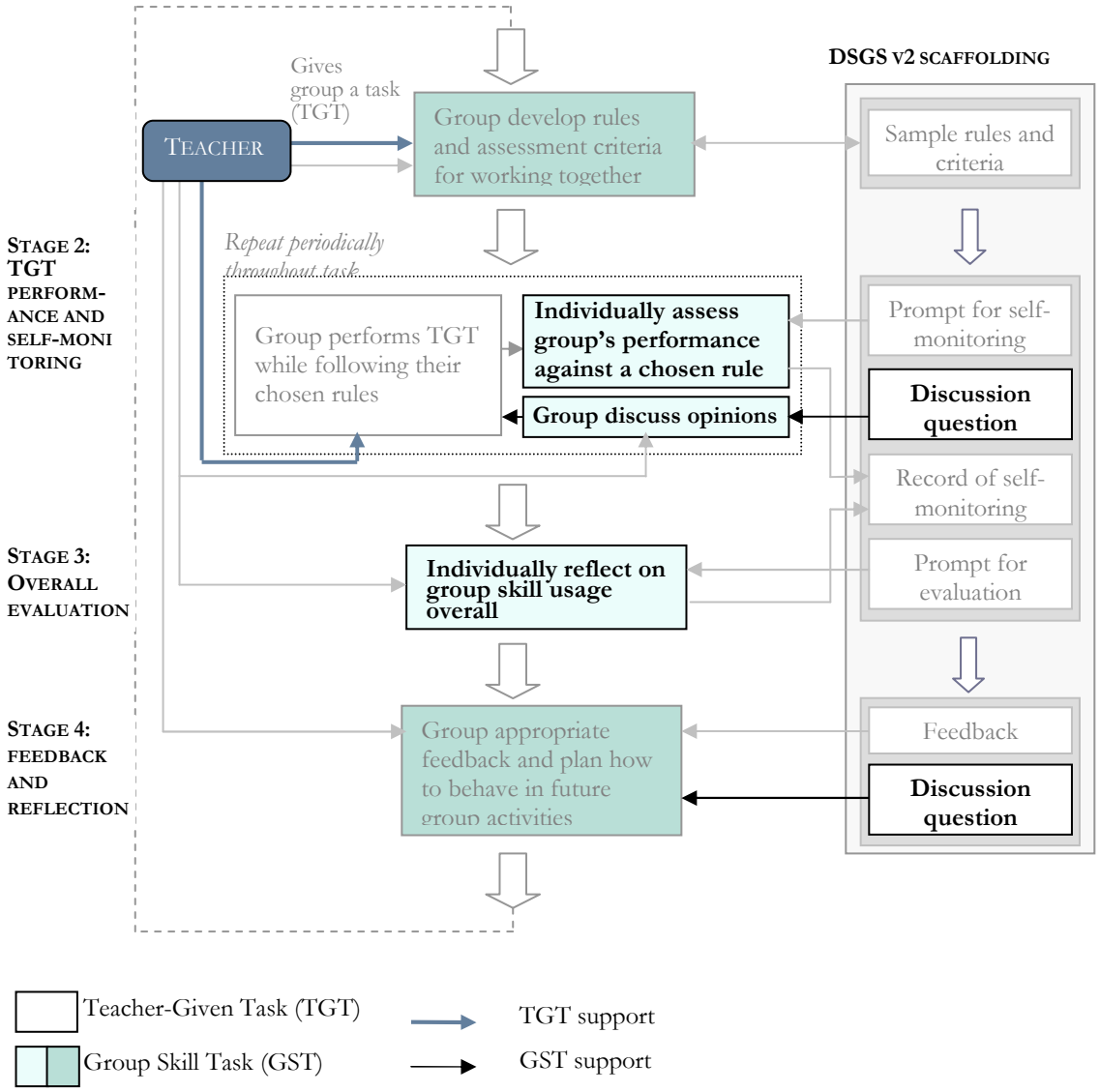


Figure 5.1: The additional stages of the DSGS v2 model overlaid in bold on the DSGS v1 model

The GST activity itself is an example of a joint activity–joint outcome task. The children are both required to self-assess, work collaboratively or collectively in answering the prompts, and the feedback is applicable to them all. In the rest of this section the theoretical bases for each of the four stages are elaborated.

5.1.1 Stage 1: Raising awareness

The effectiveness of behaviour-based interactions, as demonstrated by other researchers (e.g., Wild and Braid 1996), was evident in Study 2. It was found that selecting rules for working together and being prompted to discuss criteria for monitoring their usage led to an improvement in the TGT performance. Consequently, there is no difference in this stage to the DSGS v1 model.

5.1.2 Stage 2: TGT performance and self-monitoring

In the DSGS v1 model both children were asked to agree whether a rule had been followed and record this by clicking on it. Study 2 found, even with prompts reminding the children of the need to record group skill use, little voluntary monitoring in a classroom context occurred. To compensate for the lack of monitoring in Stage 2, the DSGS v2 model scaffolds the assessment. The steps are shown in Figure 5.1 and listed below:

1. The group is given a randomly selected rule from Stage 1 and each child independently assesses the proportion of time on a fixed 5-point scale that it was followed since the last evaluation
2. Then each child independently records this value which is stored in the "Record of self-monitoring"
3. The group answers the discussion question generated by the model

Stage 2 requires communication, as the group answers the discussion question; reflection, as the children independently make the self-assessments; and responsibility, as all must participate if the GST is to progress. This prevents "freeloaders" and one member dominating the interactions (Salomon 1995). The transfer of control to *all* the children should also overcome the detrimental effects of enforced reflection observed by Brown (1987). She observed a negative impact on the TGT performance when children were required to reflect or verbalise about their actions during the activity. In this design, the children can choose how much time is spent on reflection and the length of conversation in the GST. If no discussion occurs, however, then the DSGS v2 model reminds the children of the importance of answering the questions.

Each of the three steps is used in existing training schemes to raise group skill awareness. The use of a 5-point scale, in which children aged 9 to 11 individually assessed the proportion of time spent using a group skill after an activity, was found to be sufficient by Kagan (1992). Although in the training scheme proposed, this process of reflection is done *during* the activity to overcome the problems identified in Study 1a in accurately recalling behaviour. Children this age should be able to do the assessment task as: there is a well-defined scale for assessment, only one group skill is focused on, and the criterion was agreed in Stage 1 (Ager 1998, Finlayson and Cook 1998). The second step of recording an opinion has been used whenever completing questionnaires on group skill usage is required (e.g., Dawes *et al.* 2000).

Procedural prompting has been shown to focus discussions (e.g., Palincsar and Brown 1988, Palincsar *et al.* 1993, Slavin 1995, King *et al.* 1998). In this setting if the children record:

- All believe a group skill was *not* used then the prompt asks the children to plan how this usage could be improved
- All believe a group skill was used *all the time* the prompt requires them to justify this belief
- *Different opinions* then the prompt suggests they should redefine and agree the criteria used for assessment

The explanations offered by the group as part of these discussions can assist cognitive structuring. The children generate new understandings and re-evaluate their own belief structures to incorporate the perspectives of other group members (Säljö 1998). Having to agree vocabulary so that it is used in the same way suggests the children will use the terms as experienced members of the culture such as their teachers (Wertsch and Stone 1985).

The children's discussions are enhanced by the incorporation of scripted co-ordination, the assigning of roles to group members (Burton *et al.* 1997, Goodman *et al.* 1998 etc.). Depending on group size the model randomly allocates roles such as executor, reflector, constructor, checker and repairer, which ensures the participation necessary for learning (Baker and Lund 1997).

The prompts focus the discussion, and the roles ensure that all the members contribute ideas or examples about their behaviour, or evaluate the ideas of others. The prompted discussions should improve the TGT performance; for example, by incorporating the plan of how group skill usage could be increased within the TGT. Explicitly encouraging the children to consider their recent behaviour so that future behaviour can be amended can be considered as making reflection-in-action explicit (Schön 1987).

The second purpose of the prompts is to model the type of exchange required when working in groups (Tharp 1993). Checking for agreement, ensuring participation, and listening to each other, are skills that are required for effectively completing the TGT. Through practising such dialogues as part of the GST, they may be internalised and transferred to the TGT.

5.1.3 Stage 3: Overall evaluation

Overall group skill usage is the main focus in Stage 3. The children are *all* responsible for this stage of the GST. It requires them to individually decide and enter their own opinion of the group's use of group skills. The child is asked to reflect on the behaviour of others as this results in a more accurate assessment than asking about his own performance (Dillenbourg 1999).

It is hypothesised that this evaluation of group skill usage in the TGT will be more accurate than in earlier studies. The children should be aware of their behaviour through externalising what happened and why as part of Stage 2. This knowledge should reduce any incongruities in overall assessment (Johnson-Laird 1986).

5.1.4 Stage 4: Feedback

The feedback given to the children by the DSGS v2 model is a combination of the self-assessments from Stages 2 and 3. A numeric value is used to represent the proportion of time spent using group skill as it can be easily interpreted by children (Gallimore and Tharp 1990). This value can provide intrinsic motivation, the pride in a high mark, or, if a record is kept of group skill usage and if the scoring is consistent, the child's progress in using group skills (Ryan and Deci 2000). These marks also have an extrinsic motivational value, by a comparison with peers or getting higher values than previously. The appropriation occurs in the discussion the score provokes, not the numeric value itself. Future behaviour is influenced by the reflection and internalisation (Newman *et al.* 1989, Newman 1990).

In addition to the score, prompts are given to assist the children appropriate the feedback, that is, understand why the system gave that numeric value to represent the proportion of time spent using group skills. If, for example, the DSGS v2 model returned a value lower than their opinion the group would be asked to explain why. These prompts incorporate scripted co-ordination. One child is randomly selected to provide a plan or justification, with another asked to act as a critic (Goodman *et al.* 1998). This requirement to reflect upon one's behaviour and plan how future exchanges could be improved is a similar process to reflection-on-action (Schön 1983).

5.1.5 Summary of the DSGS v2 model

The DSGS v2 model expands the DSGS v1 model by increasing the usage of the key attributes for working well as a group: communication, responsibility and reflection. It incorporates existing training strategies as summarised in Table 5.1.

SKILLS	METHOD OF INCORPORATION
Communication	
Behaviour based interactions	• The group must agree appropriate group skills and criteria for their assessment (Stage 1)
Scripted pair learning (procedural prompting)	• Prompts based on generic question stems which use data from the children's self-assessments are used to stimulate discussion about the opinions and ways of improvement (Stages 2 and 4)
Scripted co-operation	• The roles of evaluator and critic are randomly exchanged (Stages 2 and 4)
Responsibility	• All are individually responsible for assessing the group's performance using on group skill during the task (Stage 2) • All are individually responsible for assessing the group's performance overall at the end of the task (Stage 3) • The GST feedback relates to the group performance and all are equally responsible for their actions
Reflection	• As a group deciding appropriate group skills and criteria for assessment (Stage 1) • The children reflect in the self-assessing stages, and in the discussion after the feedback is given (Stages 2, 3 and 4)

Table 5.1: Methods used to develop group skills

5.2 A computer implementation of the DSGS v2 model

This section describes one possible instantiation of the DSGS v2 model. It illustrates how the underpinning theoretical basis could be implemented using a computer as the medium. It begins by giving an overview of the interface, which was developed with children acting as informants (Druin 2002). The remainder then describes the implementation of each stage in turn.

Aliens was designed for pairs of children working at a computer. From Studies 1b and 2, the software was designed to be domain independent, since knowledge of the TGT performance is not required to assess the amount of time spent using group skills. Eight 9 and 10 year olds of varying abilities assisted in the development of the Aliens software. These children did not participate in any other study. They worked with the same partner for five one-hour tape-recorded sessions; there were two pairs of girls, one of boys, and one mixed. The teachers reviewed the software to ensure the wording was appropriate. Aliens was written in Macromedia Director v7.

The children in the design phase regarded the interface as critical, and it was therefore unsurprising that they suggested characters were required to stop the application being "boring". Anthropomorphism is motivating in younger age groups, as shown by the NIMIS project (Hoppe *et al.* 2000). Lester *et al.* (1997) found a similar effect among middle school children. They found that knowledge improves with animated characters; even without the software incorporating teaching strategies the groups appeared more motivated, gave more explanations and reflected on their answers. After nominating various creatures, the children voted for the Aliens shown in Figure 5.2. The animation, however, was restricted: both walked, one blew bubbles and the other waved. The children argued that two characters reinforced the idea of working together.

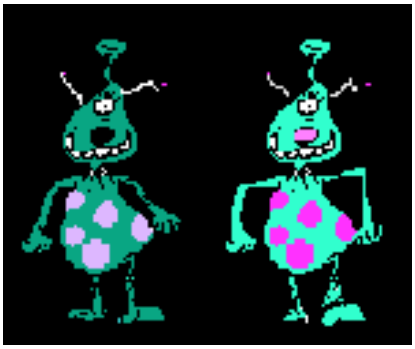


Figure 5.2: The Aliens

In addition to anthropomorphism, the interface incorporated the children's names, as shown in Figure 5.3. This appeared to be motivating, and led to the screens being spontaneously read out loud, even by the more able pupils. It may also lead to a higher recall of the information presented, as found with adults using software with a personalised style of instruction (Moreno and Mayer 2000).

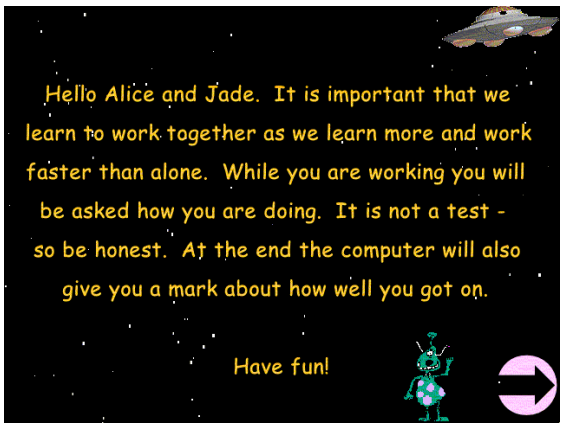


Figure 5.3: Introductory screen to Aliens

The Aliens interface has minimal typing, requiring only that the children enter their names and possibly some rules for working together, and no discrete turn taking (Wegerif and Dawes 1998, Druin 2002). The absence of typing was not considered detrimental as:

- Most communication between children is done verbally or through gestures (Cummings 1988)
- The discussions generated by prompts can be more valuable than if the children have to record their opinions on a system (Enyedy *et al.* 1997)

The software does have some degree of unilateral control, as the child with the mouse can proceed to the next question without the consent of his partner. There are also restrictions due to the computer operating system available. This meant the active application, the GST or TGT, is displayed over all other windows. This contrasts with the need to have the list of rules selected for working together being displayed throughout so they can be referred to (e.g., Dawes *et al.* 2000). Their absence may reduce the focus on the GST, as the children are only explicitly reminded to reflect on group skills when the GST interrupts the TGT.

5.2.1 Stage 1: Raising awareness

A database was constructed to provide sample rules for working together. The interface is shown in Figure 5.4. The rules it contained were:

- Agree before you perform an action
- Do not cheat
- Listen to each other
- Make sure you both understand
- Take turns

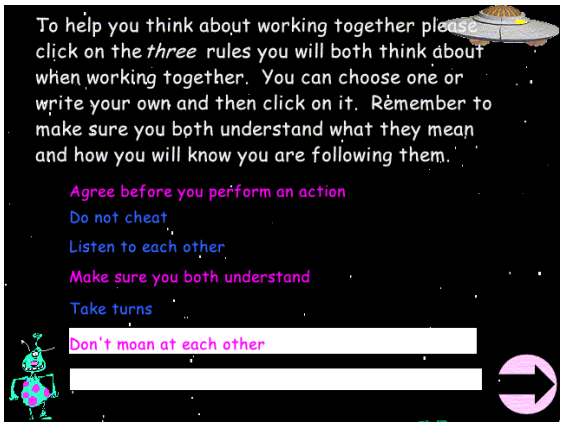


Figure 5.4: Stage 1 screenshot from Aliens

Two rules were dropped from the database used in The Escape from the Forbidden Forest software (Section 4.2.4). The first, "Do not fight", was removed because the children felt it was implicit and it had been rarely selected in Study 2. The other, "Help each other", was said to be redundant as it incorporated the rules checking for understanding, listening, and agreement. As in the DSGS v1 implementation, the children use their knowledge, teacher support and the database to select or invent three rules for working well together and criteria for their assessment. A rule could be selected by double clicking, or typed in using the existing rules as prompts.

5.2.2 Stage 2: TGT performance and self-monitoring

In this implementation the five self-assessment prompts during the TGT occur every five minutes. This is to ensure the GST can be performed within the standard 45-minute lesson and have time for TGT instructions and logging on and off. To further accommodate the context, having to finish early for example, a control sequence enables the teacher to move directly to Stage 3 if necessary.

To begin the process the child with the mouse clicks on the button to choose a rule, as shown in Figure 5.5. This is then replaced by one of their chosen rules randomly selected. The children independently select the proportion of time they believed they followed that rule since the last assessment. The five choices are shown on the screen. One child uses the mouse and clicks on the appropriate statement, the other types the corresponding number using the keyboard as instructed by the software. The time for self-assessment lasts from the click that chose the rule to the second child entering his opinion.

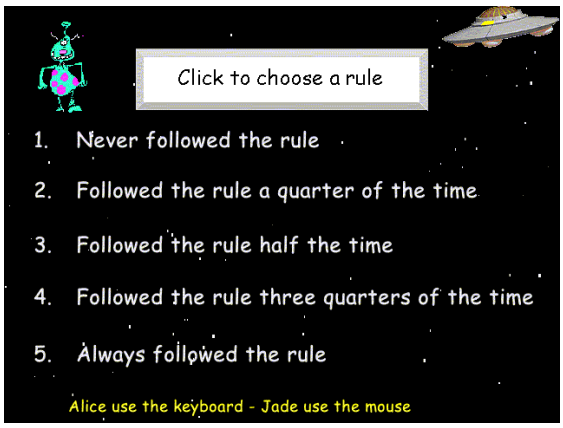


Figure 5.5: Screen prompt for individual reflection on group skills used during the TGT

Those children advising on the design recommended that the selected statement should not change colour. This would ensure they would not be influenced by their partner's selection. The children's first selection is the one stored in the "Record of self-monitoring". Once both children have entered their opinion, a discussion question is selected from Table 5.2. The scripted co-operation comes from requiring the participants to take on specific roles. Instead of the "thinkist and typist" scenario (Sheingold *et al.* 1984) one child is given an *executive* role (offering suggestions, solutions or actions) and the other acts as a *critic* (questioning or judging the responses) (Goodman *et al.* 1998).

ROLE	SELF-ASSESSMENT OPINIONS		
	BOTH OPINIONS ≥ 4	BOTH OPINIONS ≤ 2	OPINIONS DIFFER BY ≥ 2
Child - *executive*	Give an example	Suggest a plan to improve that group skill usage	Explain why you think the opinions differed
Child - *critic*	State whether you agree or disagree and why	State whether you agree or disagree and why	State why you think this explanation is correct or not

Table 5.2: Summary of discussion prompts from Stage 2

Example prompts are shown in Figure 5.6.

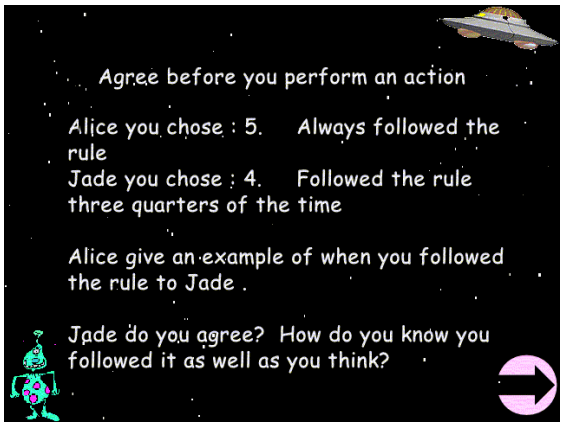
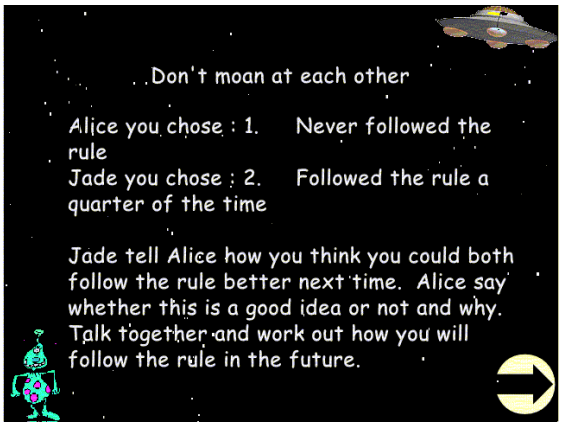


Figure 5.6a: Discussion question based on high self-assessments



b: Discussion question based on low self-assessments

Once the discussion is completed, the children click on the arrow to return to the TGT. The time difference between the appearance of the screen illustrated in Figure 5.6 and when the arrow is clicked on is recorded. This time may *not* reflect the time answering the GST question; but if the screen in Figure 5.6 is closed immediately it is assumed the prompt was ignored. If the average time the first three discussion prompts are displayed is under five seconds, one of four messages asking the children if they were sure they had talked would be displayed; an example is shown in Figure 5.7. This is to encourage discussion and the corresponding clarification of group skill definitions and criteria for assessment.

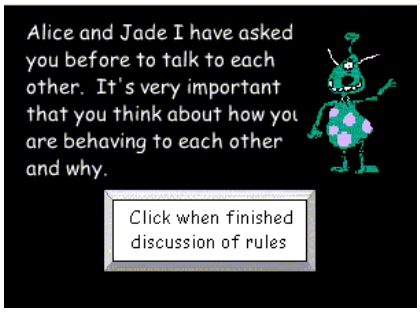


Figure 5.7: Prompt generated reminding of the need for discussion

5.2.3 Stage 3: Overall evaluation

Once Stage 2 has been repeated five times, or the instruction to jump forward has been entered, the children are asked to evaluate how well they worked together overall. Instead of agreeing a score, as in The Escape from the Forbidden Forest implementation, the dual-key control allows the children to individually enter their opinions. Their reflections are represented as a percentage: 0 – "you did not work together at all" to 100 – "you worked perfectly together". The children are allowed to change their scores until the arrow is pressed. This interface is shown in Figure 5.8.

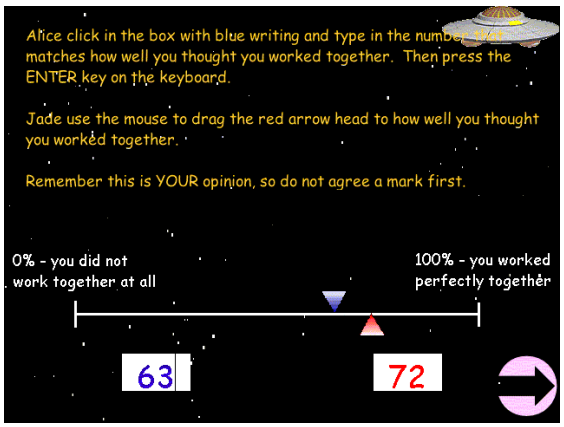


Figure 5.8: Screen for individual reflection on group skills at task end

5.2.4 Stage 4: Feedback

The Aliens software displays:

- The children's own assessment
- The software's assessment of group skill usage (the calculation of this numeric value is described in Section 7.1)
- A congratulatory message if the discrepancy between the children's average score and that of the system is less than 10%, otherwise a prompt asking the children if they can explain the difference
- Depending on the self-assessments in Stages 2 and 3 the appropriate prompt from Table 5.3 to support this reflection-on-action

AVERAGE MARK OF CHILDREN AND DSGS V2 SOFTWARE	PROMPT
Below 50	One child, randomly selected, is asked to develop a plan for improvement, the other asked to comment
Between 50 and 80	The children are asked to plan how they could improve group skill usage
Above 80	Each child is asked to give two reasons why they worked so well together

Table 5.3: Prompts generated in Stage 4

An example of the Stage 4 interface is shown in Figure 5.9.

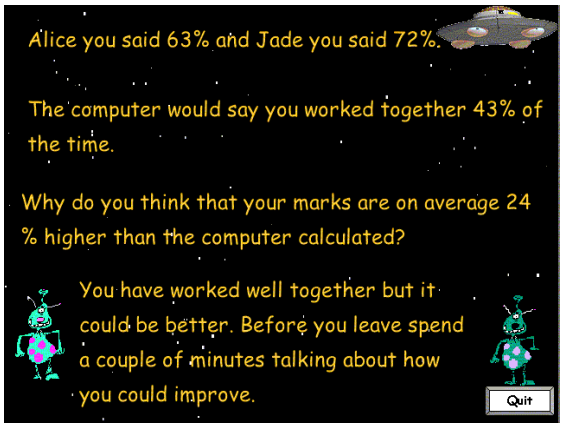


Figure 5.9: Screen with software feedback and prompts for discussion if the scores are between 50 and 80

5.3 Study 3a: The effect of compulsory reflection

In Study 1 the context was investigated so that any training scheme developed could be integrated within the classroom. These studies found the teachers focused on the objective content of the National Curriculum, i.e., topics examined in the Standard Assessment Tests (SATs) or listed explicitly - such as life in Tudor times. If a training scheme to raise awareness of group skills is to be used within these schools, it *cannot* have a negative impact on teaching this information. To test whether the compulsory self-assessment in Stage 2, during the TGT, in the DSGS v2 model was detrimental, it was partially implemented. The Escape from the Forbidden Forest exercise was adapted to incorporate the DSGS v2 model. It was hypothesised there would be no significant difference in task performance measured by the number of moves taken to solve the puzzle between this version, and that incorporating the DSGS v1 model.

5.3.1 Participants and setting

All the children participating were in Year 5, i.e., aged 9 and 10. Two schools were involved: School X, a Church of England Junior School, and School Y, a city school. The average level of achievement in mathematics, English and science in these schools was similar to the UK National Average. (See Appendix F for a comparison of the formal assessments for the schools.)

The data gathering was performed in two stages. Group A consisted of 23 pairs. They were the supported pairs from Study 2. In order to ensure the children were unaware of the TGT solution Group B were formed of children from the year below Group A but when they were in Year 5. It consisted of six pairs as the teachers having to fulfil curriculum requirements restricted access to the children. An equal number of children came from each school in each group.

5.3.2 Method

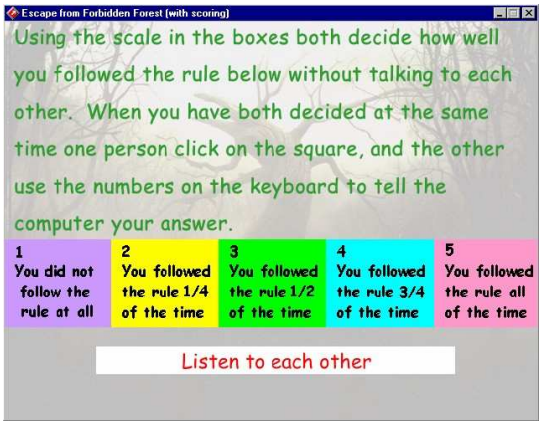
The TGT was a computer puzzle, The Escape from the Forbidden Forest (a variation of the Tower of Hanoi problem). The task was equivalent to moving four rings (the characters), between four poles (the trees). Children worked in pairs and had dual-key control; one child used the mouse to select a character, and the other the keyboard to type the number of the tree the character should climb. Initially they worked together to develop a solution. When they were confident they understood how to solve the puzzle they could choose to solve the problem alone with different starting and finishing trees. A full description of the TGT is given in Section 4.2.3.

Children in Group A performed the puzzle as part of an ICT lesson and were observed by their class teacher. Children in Group B worked in pairs with the researcher in an otherwise empty classroom. Each session took approximately 45 minutes.

Both groups began the task by agreeing three group skills in the form of rules for working together. As the TGT and GST was combined, the rules chosen were displayed throughout the puzzle and appeared in the top left corner of the screen.

Children in Group A used the variation of The Escape from the Forbidden Forest that incorporated the DSGS v1 model. If both children in the pair believed one of these rules had been followed during the TGT, they were asked to click on it. If no rule was selected after five moves, the children were reminded to think about group skill usage. When both children had successfully completed the TGT individually, the software asked them to jointly agree how well they had worked together overall and record their opinion.

The software used by children in Group B incorporated the DSGS v2 model. After five moves had been made the TGT would be paused and a new screen appear. This required them to individually assess the time spent on a randomly selected rule. This screen is shown in Figure 5.10a. Once both had answered, a discussion question was selected from Table 5.2. Unlike the Aliens implementation of the DSGS v2 model, this did not specify which child was to provide a justification. After the TGT the software asked the children to individually assess overall group skill usage. This is shown in Figure 5.10b. The dual-key control allowed for independent opinions, although a computer implementation could not ensure this occurred.



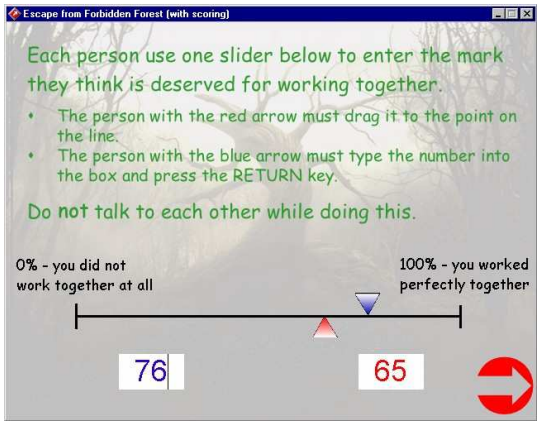


Figure 5.10a: Stage 2 self-assessment screen *b: Stage 3 overall self-assessment screen*

To corroborate the Stage 3 self-assessment all the pairs were assessed on the proportion of time group skills were used. The teachers assessed Group A, the researcher Group B. The assessment criteria used was that given by the teachers in Study 2 (see p.88). After the task, the researcher asked children from both conditions if the process of self-assessment and reflection was detrimental to their task understanding.

5.3.3 Results

There was no significant difference between the TGT performance of the children in Group A and B in: i) the first attempt to solve the puzzle, or ii) when working independently after having practised together. In Group A the average number of moves taken to first solve the problem together was 27.74, and in Group B it was 20.67 (U=64 using Mann Whitney). When tested individually children in Group A took on average 12.50 moves, and those in Group B 15.92 (U=222 using Mann Whitney).

There was no significant difference in the assessment of how well the pairs worked together between that of the teacher assessing Group A, and that of the researcher assessing children from Group B. The average mark for Group A was 58.0%, and for Group B it was 64.2%. In both groups the majority of the children's own assessments of their group skill usage were higher than the teacher's or the researcher's opinion. Children from Group B, however, were more likely to give themselves non-zero marks less than 50%, and less likely to give marks greater than 90% when assessing group skill usage. This is shown in Table 5.4. The overall assessments in Group B contrast with those recorded during the TGT. They stated that they followed the rule all the

time in 77% of the recorded statements, and three-quarters of the time in another 15% of these statements. There was no correlation between their overall assessments in Stage 3, with their self-assessments during the TGT in Stage 2 ($r_s = 0.1616$).

	TEACHER'S MEAN (%)	CHILDREN'S MEAN (%)	NUMBER OF PAIRS WHO GAVE THEMSELVES LESS THAN 50%		NUMBER OF PAIRS WHO GAVE THEMSELVES MORE THAN 90%	
Group A *Standard deviation*	58.0 *15.37*	76.7 *30.27*	3	13.0%	9	39.1%
Group B *Standard deviation*	64.2 *13.62*	72.3 *24.00*	3	25.0%	3	25.0%

Table 5.4: Comparison of group assessment scores for working together on task completion

In both groups, the children believed being asked to assess and reflect on group skill usage did not detract from their task performance. Group B required the researcher to explain to children how to enter their opinions, and give reminders to assess independently and to follow the screen prompts to discuss discrepancies or provide examples. The discussion questions did not promote much conversation.

5.3.4 Discussion

The lack of difference in the number of moves indicates the process of compulsory reflection as part of the GST is *not* detrimental to the TGT performance. Additionally, the children believe the compulsory evaluation does not impede their TGT understanding. The lack of discussion, despite the prompts, could indicate that the GST is less important to the children than the TGT. Alternatively, it may be due to the time required to process the TGT (Brown 1987). Another explanation, as suggested in Study 1b, is that many children this age cannot reflect on their own behaviour. This is in spite of:

- Research showing metacognitive skills are found in children over nine (Azmitia 1998, Sharples 1999)
- Even younger children being able to assess such skills (Bennett and Dunne 1992e)

It is possible the questions were ignored as they were not specific enough; for example, neither child was asked by name to provide a reason for their opinion.

No conclusion can be drawn about the effect of compulsory assessment and reflection on the group skill usage based on the opinions of the teacher and the researcher. The difference of approximately 5% between the average assessment given by the teacher and researcher when assessing the amount of group skills displayed may indicate:

- The teacher and researcher have a slightly different understanding of the assessment criteria
- The teacher is incorporating their knowledge of the children while assessing
- Group B's behaviour was effected by the direct observation of the researcher (Biott and Easen 1994b)

The trend for lower overall self-assessment scores by Group B, who used the DSGS v2 model, suggests compulsory self-monitoring may influence the final evaluations. It cannot be concluded these scores are more realistic since there is no correlation between Stage 2 and Stage 3 opinions. This suggests that the children are not recalling behaviour during the task when making their final assessments. This does not support the hypothesis proposed in Section 5.1.2 that answering the discussion questions would refine the children's understanding of group skills and assist in raising incongruities in later self-assessments. This inconsistency between the assessments may be related to the lack of discussion and externalisation in Stage 2 necessary for refinement of understanding (Vygotsky 1978).

The study raised one design issue. To be effective the introductory screens had to be reiterated by the researcher. Stressing the importance of following the prompts in a written format did not result in the children reading and reflecting on them. Without this support the children appear to focus on the TGT and *not* the GST. This supports the idea that children need to know the goal of the task (Lin 2001, Kramarski and Mevarech 1997, Salomon 1984), and suggests a written statement is insufficient to cause this understanding.

5.4 Study 3b: The effect on monitoring

The feedback generated by the DSGS v2 model and given to the children in Stage 4 is dependent on the accuracy of the information entered. Studies 1, 2 and 3a all showed that compared to an adult children overestimate group skill usage when considering behaviour retrospectively. Despite self-assessment being advocated (Ager 1998) children are not familiar with this process (Weeden *et al.* 1999, Study 1b, 1d). Accuracy in self-assessments during the task in the DSGS v2 model is improved by requiring the children to:

- Consider *one* aspect of their behaviour using criteria agreed between them
- Assess this usage on a well-defined scale

Study 3a showed that children aged 9 and 10 believed they followed the rules for working together all the time in the majority of self-assessments during the activity. To determine whether such high marks are due to an inability to relate their actions to their understanding - as suggested in Study 1a, or children are incapable of recognising group skill usage groups were asked to assess three scenarios. This would:

- Identify whether children can distinguish between different types of group exchanges
- Determine if children can focus on one group skill and assess the proportion of time it is used
- Ascertain the impact of the training scheme on the self-assessments by comparing the opinions of general group skill usage from children using the DSGS v2 model with those who are not

5.4.1 Participants and setting

Two Year 5 classes from School X and School Y participated. The teachers ranked their pupils by ability. Two matched groups were formed in each school by taking all the children assigned an odd number from one class and combining with those given even numbers from the other. Group C was the control group; Group D used the DSGS v2 model. Each group was subdivided into sets of similar ability according to the teacher's ranking and as far as possible contained four children from the same class. Due to absences, 51 children from School X took part, six sets that used the DSGS v2 software, and five sets in the control group; and 46 children from School Y took part, six sets for each group, however one of the control groups was excluded due to misbehaviour.

The children had just completed Study 1e. The activity was held in the library, or in an otherwise empty classroom.

5.4.2 Method

Three scenarios in the form of short scripts were used. These had been developed with six Year 5 children not participating in the study. All of the scenarios involve two characters performing a spreadsheet exercise. In the first the characters fight over whose turn it is to do the typing, in the second they agree with each other but work independently, and in the third, they try to work out the task and complete it together. In Mercer's terms, the style of exchange would be classified as disputational, cumulative and exploratory (1995b, p.104). The children assisting in the design confirmed the language was suitable and agreed the scripts were ordered according to the amount of group work displayed. The scripts are in Box 5.1.

Alex and Sam are at the computer. They are about to work on a spreadsheet.
Instructions for the actions taken by them are given in brackets.

Scenario 1

Alex: What are we suppose to do?
Sam: I don't know stupid!
Alex: Stupid yourself - you've got the questions - give them to me.
Sam: NO… It says here we've got to type in a list of numbers and then get the computer to add them up…
Alex: I'll do the typing
Sam (interrupting): No I'll do it - you did it last time.
Alex: No I never!
Sam: Yes you did!
Alex: Didn't!
Sam: Did!

Scenario 2

Alex: What are we suppose to be doing?
Sam: We've got to type some numbers in and then get the computer to add them up using a formula.
Alex: OK then, you do the first one and I'll do the second.
Sam: Can you pass me the mouse then and I'll call you when I'm done.

Scenario 3

Alex: The instructions say we have to type in a list of numbers and then get the computer to add them up using a formula.
Sam: What's a formula?
Alex: It says here that it's an instruction to the computer - rather like a sum in our maths book.
Sam: I still don't understand.
Alex: (Holds sheet in front of them) Yes, I see… look… this says one formula could be A1+A2, which is the number in cell A1 - here (points to the screen) added to the number in the cell below (points to it). The formula is the bit that tells you what to add.
Sam: So if I wanted to add what's in that cell and that cell I would put C3+C4?
Alex: Yeah I think so.
Sam: So what we've got to do is type in the numbers here, then type in a formula that will add them all up, that goes in the cell under the numbers, and then we're finished.
Alex: Yeah… you read me the first 2 numbers and I'll type, and then we can swap for the next two.
Sam: OK then.

Box 5.1: Scenarios for assessment exercise

The children were told that as part of the National Curriculum the teachers were required to assess how well they worked together. This made the task more realistic and provided a context (Biott and Easen 1994c). The children were given two copies of one of the scripts and asked to negotiate who should read aloud and then act out the scenario. Then they had to assess how well they thought the children had worked together. Once this was done, the next scenario was given; the sequence was Scenario 1, 2 and 3. Once all the scenarios had been assessed, the children were given all the scripts and told to rank them in order of increasing group work displayed. Finally, the children were asked to assess their own performance as a group. It was designed to be a joint activity–joint outcome task as the children were meant to work together to reach a shared conclusion. The exercise was tape-recorded and lasted about 15-20 minutes.

When assessing the scenarios the children in Group C were asked to pretend they were teachers. They were given a marking scheme: 0 meant the children in the scenario did not work together at all, and 100 meant they worked perfectly together. First, they individually decided a mark about the how well the pair had worked together and wrote it down. Then as a group, they were told to agree a mark and provide justification for it. This was also written on their sheets. The children were reminded to consider the rules they had discussed in Study 1e, for example, the importance of helping each other, sharing ideas, being nice, and listening. If no justifications were provided

the researcher would prompt for reasons. The children were told there was no right or wrong answer.

Those sets in Group D used the DSGS v2 model in the form of the Aliens software to assess the scenarios. Before being given the first script they were asked to select three rules, either their own or from the database given in Section 5.2.1. Then they discussed and agreed the criteria for assessing whether each rule had been followed. The researcher would prompt them if no discussion occurred. Next, the children were asked to divide into two groups of roughly equal size. To assess a scenario each subgroup had to decide how well the children in the scenario followed *one* of their rules randomly selected by the software. They used a scale of 1 – not at all, to 5 – followed it perfectly. Once each subgroup had discussed and agreed a score, one half used the keyboard to type in the appropriate number, the other the mouse to click on the appropriate level. As described in Table 5.2, according to the extent to which the opinions agreed or disagreed, one subgroup was prompted to: i) justify their answer, ii) explain the difference in opinion, or iii) suggest methods by which the pair could achieve a better mark. Once this was completed, they were given the next script. The researcher stressed the importance of this discussion stage and would prompt if the questions were not answered.

To assess their group's usage of group skills, those in Group C were asked to individually decide and record the proportion of time their set had spent using group skills. They were then asked to discuss these values and agree a mark as a group, providing a justification for this figure. Those in Group D completed Stages 3 and 4 of the DSGS v2 model. That is, in their subgroups they reflected on group skill usage and determined the proportion of time group skills had been used then entered this value into the computer. The Aliens software generated a numeric value and prompted them to explain any discrepancies, plan future group skill use, or give examples justifying their opinion. The paper records made by Group C and computer log files generated by Group D were analysed.

5.4.3 Results

The children had no difficulty providing a score relating to the level of group skill usage in the scenarios. Providing justification for these opinions initially required prompting, as shown in the first transcript extract below. Further prompting was needed initially with Group D. Instead of focusing on the selected rule, they would consider how the characters worked together overall.

To justify why they had said the characters followed the rule "help each other" three-quarters of the time one set from Group D gave not fighting and listening as reasons. As the task continued, those in Group D began to justify their reasoning and answer the questions given by the Aliens software without reminders; this is illustrated in the second transcript extract.

1. *Discussion of Scenario 2 within a set from Group C:*

Pupil 1: About 75
Pupil 2: Yeah 75
Pupil 3: 75-80
Pupil 4: Yeah 80
Pupil 3: 80
Silence - researcher prompts for a justification.
Pupil 3: Because like…
Pupil 4: Because they were like working together and they never ever argued <*quiet chattering from other girls in background*>
Pupil 2: There were no arguments and they said when I finish then it's your turn.
Pupil 3: Yeah - Like sharing.
Pupil 1: {they got on quite well}
Pupil 2: {they shared the thing}
Pupil 3: They got on alright but like they should have said please and stuff.
Pupil 4: But they didn't say stuuupid!

2. *Discussion deciding score of children in Scenario 2 using the rule "Take turns" by a pair of children in Group D:*

Pupil 1: Five because when they were reading it out they were like sharing it more…
Pupil 2: like saying when he said I'll do the first one and you do the second and Alex said ok then you do the first one and I'll do the second.
Pupil 1: They were talking after each other and weren't interrupting.

Both groups asked for confirmation about the opinions given in the self-assessment process.

All 21 groups that completed the exercise agreed that the first scenario had the least group work, as they were arguing in a destructive manner. The given marks ranged from 0 to 35 in Group C, and 1 to 4 in Group D. The scores for the second and third scenarios were high. Only one set in ten from Group C gave a score of less than 95 on first reading for Scenarios 2 and 3. Three sets from 11 in Group D did not give a perfect score for group skill usage in Scenario 2. This fell to two in Scenario 3.

The first three data columns in Table 5.5 show the ordering after first reading. Half the Group C sets believed Scenario 3 showed the most group work. Over twice as many sets in Group D

compared to Group C judged the second and third scripts to display an equal amount of group work on the first reading. There were no significant differences between the assessments made when the children were asked to assess the whole exchange, as in Group C; and those made were the children were asked to focus on one rule, as in Group D.

The "Assessment after reading all" columns in Table 5.5 show the number of groups choosing that order of scenarios when asked to rank them at the end of the activity. The number that believed Scenario 3 showed the most group work increased for Groups C and D. Despite being told to consider group skill usage overall Group D still has a higher proportion of sets who believed Scenarios 2 and 3 showed the same amount of group work. The reason given for the third script being slightly better by the majority of children when ranking the scenarios was that it had more explaining. This was summarised by the statement: "That one's [scenario 2] just working together, but that one's [scenario 3] helping and working together."

The bottom half of Table 5.5 shows the same information grouped by the schools. Compared to School X a significant number of children from School Y after reading all the scripts ranked the cumulative style of working to be better group work than the exploratory approach ($p>0.05$ using Chi square). Reasons given by the four groups include: "they had to ask", "I think – I think he worked well, but he didn't understand the formula", "number 2 is better as they start working sooner".

Children's ordering of scenarios by increasing group skills usage	ASSESSMENT ON FIRST READING						ASSESSMENT AFTER READING ALL					
	1, 2, 3		1, 3, 2		1, 2=3 (2 and 3 ranked the same)		1, 2, 3		1, 3, 2		1, 2=3 (2 and 3 ranked the same)	
		%		*%*		*%*		*%*		*%*		*%*
Group C (10 groups - 2 misbehaved)	5	*50.00*	3	*30.00*	2	*20.00*	7	*70.00*	2	*20.00*	1	*10.00*
Group D (11 groups)	4	*36.36*	2	*18.18*	5	*45.45*	5	*45.45*	2	*18.18*	4	*36.36*
School X (11 groups)	4	*36.36*	2	*18.18*	5	*45.45*	7	*63.64*	0	*0.00*	4	*36.36*
School Y (10 groups - 2 misbehaved)	4	*40.00*	3	*30.00*	3	*30.00*	5	*50.00*	4	*40.00*	1	*10.00*

Table 5.5 Comparison of script assessment

There was a wide range in the overall self-assessments. Group C averaged 73.36% (standard deviation 33.12), and Group D 76.29% (standard deviation 31.52). The children considered their use of general learning skills when assessing how they worked together. They deducted marks for speaking simultaneously, for arguing, and for proceeding with the task without checking that everyone agreed. Over half the groups used the scenarios to justify their opinions, for example, saying they deserved a lower score than Scenario 3 as they offered fewer explanations.

5.4.4 Discussion

The task of assessing scenarios is not representative of assessing group skill usage as part of the GST while performing a TGT. This is because the TGT *was* to assess group skill usage. Other reasons that make this assessment uncharacteristic are the children: i) are reviewing group skill usage over a shorter time, ii) can refer to the scripts, and, iii) are being directly observed. This combination of the TGT with the GST, however, avoids the difficulty that children would have in trying to focus on two activities concurrently (Brown 1987). By combining the tasks the results indicate it is possible for the children to:

- Distinguish between the levels of group skills used
- Focus on one rule for effective group work
- Incorporate these opinions in later assessments

The scores show all the children can recognise disputational talk and do not classify it as working together. It is more problematic to distinguish between the amount of time group skills are used if they exist, as in cumulative and exploratory discussions. This is demonstrated by the high assessments for Scenarios 2 and 3 by Group C. The same values by Group D could be due to difficulties in recognition, or could demonstrate that they are capable of considering individual rules. As the second transcript on page 122 shows the children can validate their reasons for why the pair in Scenario 2 took turns, yet overall the group skills demonstrated in this scenario are less than those in Scenario 3. These discussions could explain why more children in Group D recorded Scenarios 2 and 3 as using an equal amount of group work. The majority in Group C could distinguish between these scenarios, those 20% that did not may suggest that the pervasive turn taking culture described by Crook (1998) is prevalent. These children appear to judge success of the group by the characters sharing the task, rather than co-constructing a solution.

The fact that some in Group D could answer the prompts and provide justifications, an example of exploratory talk, suggests that they can focus on one specific rule although prompts by the researcher assist discussion. Furthermore, as shown in the second transcript this focus leads to a clarification of that skill's description, rather than a general discussion. In this case, the computer appeared to be acting as a mediating artefact; the environment it provided allowing the group to discuss and develop an understanding of group skills. However, these discussions seemed to influence the children's opinion of a scenario. Thus, when asked to compare the scenarios, some of those in Group D maintain their belief that Scenario 2 and 3 display equal amounts of group skills: while others were able to recognise the incongruity between their opinions, which led to a reassessment. This implies that future implementations of the model need to stress the difference between assessing one rule and overall behaviour – which takes into account other group skills.

The self-assessment overall was not impacted by separating the GST from the TGT. As suggested in Study 1b, knowledge of general learning skills usage was sufficient for the children to evaluate the amount of group work. The assessment of the scenarios, however, did impact the children's beliefs about their own performance. The comparison with the scenarios would explain the lower average mark for the self-assessments in Groups C and D, compared to their assessments of Scenarios 2 and 3. This concurs with Johnson-Laird (1986) – that if an incongruity is raised, i.e., if the children have a concept of effective group working, they are more likely to recognise flaws in their own behaviour.

The assessment process is strongly influenced by the school. The significant number of children that believed Scenario 2 displayed most group skills in School Y suggests they used a different set of criteria to School X. As shown in Study 1d, teaching in both schools is didactic, but in School Y the immaturity of the pupils results in a great deal of time spent on classroom management. This was reflected by the reviews of videos of children working together in Study 2 – teachers in School Y penalised noise. School Y appears to have a hidden curriculum in which the children perceive quiet behaviour and turn taking as required attributes for effective group working (Branco 2001). Cumulative talk, with its focus on agreement, fulfils this requirement, unlike exploratory talk, which may include arguments even though they might be of a constructive nature (Mercer 1994). This results reinforce the idea that the outcome is influenced by the context (e.g., Forman and Cazden 1985).

The number of prompts for justification, large numbers who believe the characters in Scenarios 2 and 3 work perfectly, high overall self-assessments, and constant checking of opinion validity support the claim that the children are not used to assessing themselves or others (Weeden *et al.* 1999). The debate stimulated by the scenarios, however, supports the claim that children find it easier to consider the actions of others (Dillenbourg 1999).

5.5 Summary

In this chapter, the refinements suggested by Study 2 to overcome the lack of self-monitoring in the DSGS v1 model were discussed. In addition to compulsory self-assessment, considering the proportion of time spent using a group skill, and requiring individual opinions, a further enhancement, the addition of prompts, has been described. These prompts, which occur during the activity and in the appropriation of feedback, are based on procedural prompting. They are designed to explicitly raise awareness of interactions and to model the type of talk expected within the TGT. They require the children to provided justifications, plans, and examples. Through the incorporation of scripted co-operation, participation is required. Like DSGS v1, the DSGS v2 model still uses behaviour-based interactions to raise awareness of group skills before commencing the TGT. Responsibility is enhanced through making each group member responsible for entering their own opinion on group skill usage, and if necessary having to justify their view.

The DSGS v2 model was implemented as a software application for pairs called Aliens. This is domain independent. It was designed with the help of children and motivation is enhanced by anthropomorphism and personalised instruction.

To test whether compulsory reflection was detrimental to TGT understanding, the number of moves taken to solve a task when using the DSGS v1 model was compared to that needed when using the DSGS v2 model (Study 3a). No significant impact on performance was found. The prompts for discussion, however, did not generate much talk in the DSGS v2 setting. This may explain the lack of correlation between the self-assessments during the activity and overall when using the DSGS v2 model.

Study 3b focused on the assessment of group skill usage. Studies 1and 2 showed that children overestimate overall group skill usage compared to a teacher's or a researcher's opinion. Study 3a

had shown the children believed that during the task they used group skills "all the time" for the majority of self-assessments. As feedback is dependent on the accuracy of the information recorded, this study investigated whether it was the vocabulary, or recall of actions, that was leading to the high self-assessments.

Using scenarios suggests that children can recognise disputational talk, but have difficulty in distinguishing cumulative and exploratory talk. Both are considered as using high amounts of group skills. Despite this similarity, the majority of children can rank the scenarios by group skill usage. Furthermore, the study suggests that if children consider one group skill, as in the DSGS v2 model, the discussions generated can lead to a clarification of that skill. However, some children require more training to distinguish between being asked to assess overall group skill usage and one particular skill. The process of assessing a single group skill in the DSGS v2 model appears to influence later evaluations. If a group is believed to be following *one* rule, this influences the assessment of general group skill usage. The study also shows that an adult is required to ensure discussion occurs, even if they do not influence the talk itself. It concludes by observing the greatest impact on the recorded assessments is not whether the children are asked to focus on overall group skill usage or one specific rule, but the school to which the children belong. The community and existing understanding appears more important than the mediating artefact in the form of the training scheme.

Chapter 6

INVESTIGATING THE MEDIUM

Chapter 4 suggested although computers encourage talk and children find them motivating a drawback to a computer as a medium for the Developing and Supporting Group Skills (DSGS) model would be the perception that children had about information they had to enter. Study 1d suggested that children perceive computers as: i) "drill and skill" packages, ii) an information resource via the Internet or CDs, or iii) a set of skills, such as being able to make text bold in word processing. These require "correct" actions by the children. To raise awareness of group skills it is important that children record their self-assessments accurately and are not influenced by the medium, as this information is used to generate the feedback (see Section 7.1). Study 3b indicates this is possible, as the majority of children managed to assess and rank scenarios according to the amount of group skills displayed regardless of medium.

In this chapter the impact of using a computer as a medium for the DSGS v2 model on the scores recorded is investigated further. The self-assessments are contrasted with the same model implemented using pen and paper. It investigates the hypothesis within the DSGS v2 model that the continuous process of self-assessments in Stage 2 will result in an overall assessment in Stage 3 that reflects what occurred (Section 5.1.3). This was not supported by Study 3a.

6.1 Study 4a: The effect of the medium

In this study the self-assessments of two groups using a partial implementation of the DSGS v2 model, Stages 1 to 3, are compared. One group was given the instructions and prompts on paper and used paper to record their self-assessments. The other received all the instructions and prompts via a computer, and used this to record their self-assessments. The children worked outside a classroom in a laboratory-style setting.

6.1.1 Participants and setting

All the children participating were in Year 5. They came from School X, a Church of England Junior School, and School Y, a city school. When combined, the average level of achievement in mathematics, English and science in these schools was similar to the English National Average (see Appendix F for a comparison of the formal assessments for the schools).

Forty-two children participated. In the previous term they had taken part in Study 2. They were split into two groups: Group E consisted of 15 pairs, six from School X and nine from School Y, and Group F consisted of six pairs, two from School X and four from School Y. Each pair worked with the researcher in an otherwise empty classroom.

6.1.2 Method

All the children completed a spreadsheet task in pairs as the teacher-given task (TGT). The children were both responsible for completing the task but individual understanding was not tested. In addition they concurrently performed Stages 1 to 3 of the group skills task (GST) supported by the DSGS v2 model. This was implemented as a separate software application. Group E used the paper-based implementation. They were presented with a choice of rules, discussion questions etc., on printed cards, and used pen and paper to record their opinions. Group F used software incorporating the model with dual-key control for entering their opinions: one child using the mouse and the other the keyboard. The sessions lasted for 45 minutes.

Initially all the children selected or invented three rules necessary for working well together. The sample rules can be found in Section 4.2.4. The children were asked to agree criteria for assessment. This was Stage 1 of the DSGS v2 model. They were then allowed to start the spreadsheet task. For Stage 2 a rule was selected after each subtask, for example, inputting a list of sweet prices and the quantity required then working out a formula to calculate the amount this would cost. A random number generator was used for this selection by the software, and in the paper-based implementation it was selected according to a list of random numbers. As described in Section 5.2.2, a discussion question was then presented based on the responses of the children, e.g., if both selected a value of two or below the children were asked to plan how performance could be improved. The children in Group E had to show their answers to the researcher who then selected a card with the appropriate question, while the question automatically appeared on screen for those in Group F. This process was repeated approximately five times. At the end of the session, Stage 3, the children in Group E independently wrote down how well they believed they had worked together overall, while those in Group F used the dual key control to record their opinions on the computer.

While the children were performing the TGT, Stage 2, the researcher assessed how well the group worked together using the criteria provided by the teachers in Study 2. After Stage 3, the children were asked whether the self-assessment and the reflection prompts impeded their performance of the spreadsheet task.

Each session was tape-recorded. The paper score sheets created by Group E and the log files generated by Group F were used in the analysis.

6.1.3 Results

All the groups completed the spreadsheet task with similar levels of practical assistance from the researcher. None of the children considered that self-assessment or being given prompts to reflect on their behaviour was detrimental to their understanding of the task.

As shown in Table 6.1 at least 80% of the children in both scenarios chose 4 or 5 when asked how much of the time they had followed the chosen rule in Stage 2, i.e., they believed they followed the rule over three-quarters of the time. Children from both groups rarely believed they did *not* use group skills. Only four children from Group E felt they never followed a rule for working together at all in one period of assessment.

STAGE 2 SELF-ASSESSMENT	**GROUP E - PEN AND PAPER**		**GROUP F - COMPUTER**	
	Frequency	*Percent*	Frequency	*Percent*
5 (all the time)	68	*61.82*	40	*76.92*
4	21	*19.09*	8	*15.38*
3	13	*11.82*	2	*3.85*
2	4	*3.64*	2	*3.85*
1 (none of the time)	4	*3.64*	0	*0.00*

Table 6.1: Summary of self-assessment scores from Stage 2

Without prompting children in Group E frequently hid their paper from their partner whilst writing. In contrast, in approximately two thirds of cases children in Group F conferred before selecting an opinion – this was largely in concise speech or gestures, for example, pointing to the screen. Children who used paper gave a different opinion to that of their partner in 32.73 % of all self-assessments, while those who used a computer differed in 7.69% of cases (see Table 6.2). This difference in opinion about the amount of time spent following a rule for working well together is significant ($p>0.01$ using Chi square). For both groups the length of the discussions

prompted by the questions during the TGT was less than a minute. However, when sharing their answers children who used paper often spontaneously argue before being prompted about why their partner was incorrect – which involved justifying their own opinion and giving examples.

DIFF IN STAGE 2 SELF-ASSESSMENTS	GROUP E - PEN AND PAPER		GROUP F - COMPUTER	
	Frequency	*Percent*	Frequency	*Percent*
0	37	*67.27*	48	*92.31*
1	14	*25.45*	4	*7.69*
2	3	*5.45*	0	*0.00*
3	1	*1.82*	0	*0.00*
4	0	*0.00*	0	*0.00*

Table 6.2: Summary of difference in self-assessment scores between children in Stage 2

In Stage 3 there was no difference in the assessment of overall group skill usage between pairs in Group F, while four out of the 15 pairs differed in opinion in Group E. Despite the medium, both groups gave a higher mean self-assessment on group skill usage than compared to the researcher's opinion. This is shown in Table 6.3.

AVERAGE ASSESSMENT FOR GROUP SKILL USAGE OVERALL	GROUP E – PEN AND PAPER (%)	GROUP F – COMPUTER (%)
Researcher *Standard deviation*	65 *17.02*	67 *5.77*
Self-assessment in Stage 3 *Standard deviation*	85 *24.59*	84 *19.17*
Percentage of children whose Stage 3 self-assessment was greater than 90%	66.7	50.0

Table 6.3: Comparison of assessment in Stage 3, group skill usage overall

There is a strong correlation between the overall assessment performed in Stage 2 and the assessment in Stage 3 during the task in Group E ($R = 0.7407$, $p>0.001$ using Pearson product moment correlation). This relationship is not found in Group F's results, as shown in Figure 6.1.

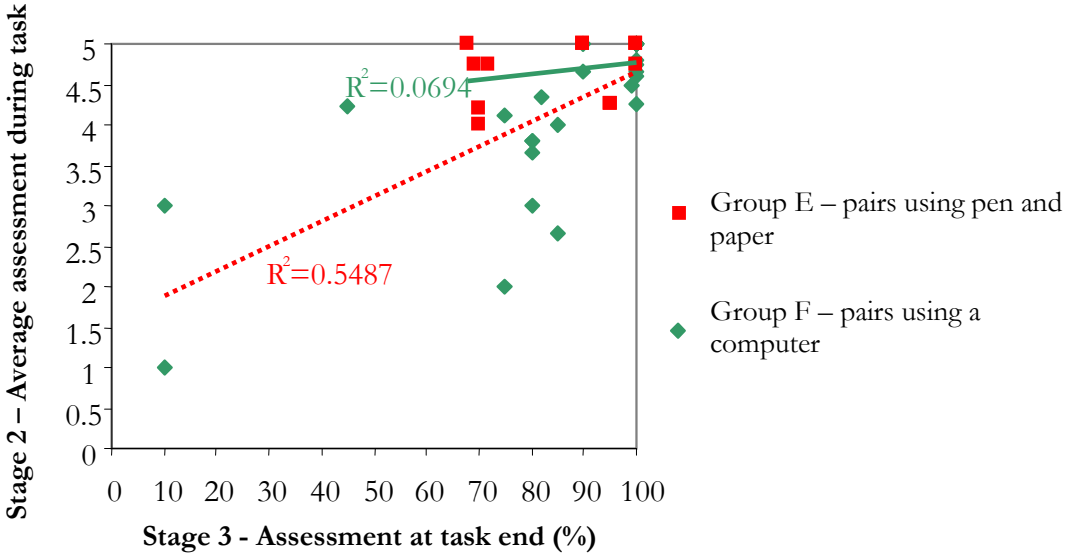


Figure 6.1: Graph showing the correlation between self-assessment during and at the end of the task in Study 4a

6.1.4 Discussion

As in Studies 1, 2 and 3, the results in Table 6.1 and Table 6.3 show that children have a high estimation of their use of group skills during the task and overall. The discussions show they believe that the process of reflection does not impede their TGT performance. This is supported by the fact the majority of time was spent on the TGT, i.e., completing the spreadsheet tasks.

Although the medium does not impact the TGT performance, it does appear to affect the children's opinions. The high level of consultation and significantly more identical self-assessments in Stage 2 between the children in Group F indicates they are not working independently. This may be due to the hardware used and preconceptions about computer applications. Having one screen enables the children to observe the opinion of their partner, and having to input an answer could be equated to a "drill and skill" package where one answer is required despite the dual-key control. If the children were working at two terminals it is possible the same amount of differences in opinion would be observed as those recording their judgements on paper.

The use of a computer also appears to impact the relationship *between* the self-assessments. In Stage 2, children in Group F are more likely to state they follow the rules all the time than those in Group E, yet overall more believe they worked together less than 90% of the time. While the opinions in Stage 2 from those in Group E correlate to their overall opinion in Stage 3 (Figure

6.1). This suggests the process of using paper, or working independently, to record opinions may raise awareness of behaviour. Paper is non-judgemental, i.e., the children believe that their opinions are valid. If there are differences in opinions, even without prompt questions, the children spontaneously use exploratory talk to clarify the definition and assessment criteria of rules for working together. By having discussions they may become more aware of their actions, enabling them to recognise incongruities in self-assessment when looking at behaviour overall (Johnson-Laird 1986). This awareness of behaviour would explain the correlation between the self-assessments observed in Group E pairs. This correlation does not suggest the children are applying criteria in the same fashion as an adult (Wertsch and Stone 1985). If this were true, the Stage 3 assessments would be similar to those of the researcher.

The findings from this study suggest that more emphasis is to be placed on independently entering opinions when using a computer and answering the discussion questions. This would enable the children to benefit from discussing discrepancies in opinion and thus developing their shared understanding of group skills and criteria for assessment. It also appears that some adult monitoring or feedback is given so that children realise they are overestimating group skill usage.

6.2 Study 4b: Corroboration of the effect of the medium

Study 4a suggests that using pen and paper to record self-assessments, rather than a single terminal, makes children more aware of their actions. Study 4b aimed to strengthen this finding of consistency between Stage 2 and 3 self-assessments when paper was used as a medium, and whether experience of the DSGS v2 model improved self-assessment consistency. It was also designed to compare the impact of the medium in a classroom setting rather than a laboratory-style study. To achieve this two sets of comparisons were made:

- Between the self-assessments recorded on paper of children performing the same TGT when one group has been exposed to the DSGS v2 model for 18 weeks and the others are using the model for the first time
- Between the self-assessment recorded on a computer, and in the following week on paper, by a group of children who have been exposed to the DSGS v2 model for 18 weeks

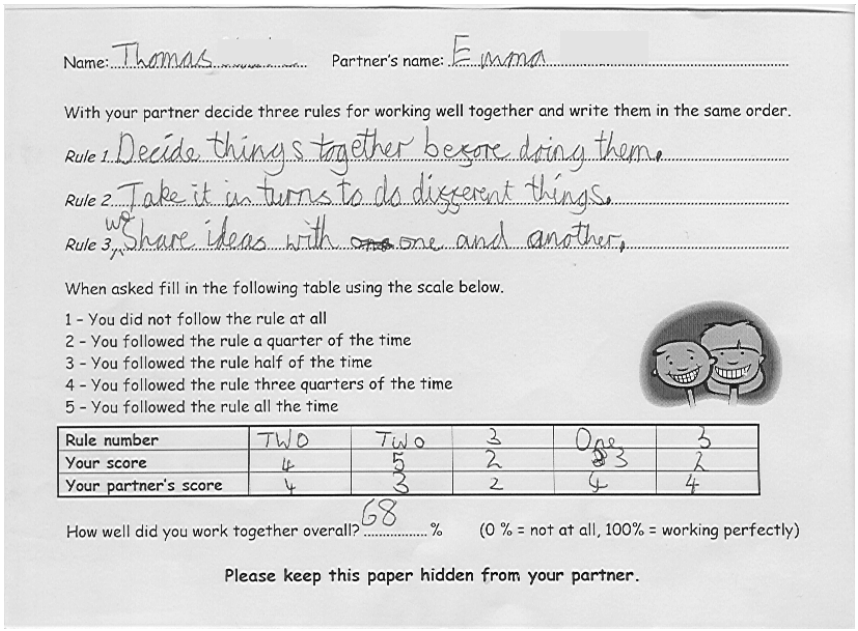
6.2.1 Participants and setting

The 48 participants came from Year 5 in School X. They were from the year below those in Study 4a. Group C, the control group, consisted of 10 pairs. There were 14 pairs in Group D, who had already used the Aliens software based on the DSGS v2 model for 18 lessons. Data was gathered in two Information and Communication Technology (ICT) lessons on consecutive weeks.

6.2.2 Method

The children worked in pairs within their 45-minute ICT lesson to complete the TGT – programming a robot on screen in week one, and looking at input and output in the next lesson (see Appendix J for both lesson plans). These lessons were based on the Department for Education and Employment's Scheme of Work (DfES 2001). The children were required to work together but unlike The Escape from the Forbidden Forest there was no dual-key control or testing of individual understanding within the TGT. Log files from these tasks indicate a similar performance between the pairs. All the lessons were a combination of the TGT and the GST supported by the DSGS v2 model.

In the first week of the study, the self-assessment values recorded by Group D for Stages 2 and 3 were taken from the log files created by the Aliens software. In the following week, the self-assessments were recorded on the form shown in Figure 6.2. This session was identical for Groups C and D. It began with the researcher asking the children to agree three rules for working together and the criteria for assessment with their partner (Stage 1). These rules were to be written in the same order as their partner on the form. For Stage 2 the researcher stopped the task at approximately five-minute intervals. She asked the children to write the randomly selected number (1, 2 or 3) in the top row of the first blank column. Then to decide *alone* how much of the time they had followed the rule with that number on their sheet since they were last stopped. Once done to write the number matching the proportion of time in the box below the rule number. Once everyone in the class had done this, they were told to write down the score of their partner in the box below. The questions from Table 5.2 were written on the room's whiteboard, and the children had to work out the difference between their scores and answer the appropriate question. For Stage 3 at the end of the lesson they were asked to individually decide and record on the sheet how well they had worked as a group overall. The sheets were then collected. (The rules chosen can be found in Appendix G4.)



Name: Thomas Partner's name: Emma

With your partner decide three rules for working well together and write them in the same order.

Rule 1 Decide things together before doing them.

Rule 2 Take it in turns to do different things.

Rule 3 We Share ideas with ~~one~~ one and another.

When asked fill in the following table using the scale below.

1 - You did not follow the rule at all
2 - You followed the rule a quarter of the time
3 - You followed the rule half of the time
4 - You followed the rule three quarters of the time
5 - You followed the rule all the time

Rule number	TWO	TWO	3	One 3	3
Your score	4	5 3	2		2
Your partner's score	4	3	2	4	4

How well did you work together overall? 68 % (0 % = not at all, 100% = working perfectly)

Please keep this paper hidden from your partner.

Figure 6.2: An example of the record sheet used with paper based implementation of the DSGS v2 model

6.2.3 Results

As Table 6.4 shows, regardless of medium over 75% of the children believed they use the group skills over three-quarters of the time. Group D self-assess themselves as using group skills for a significantly greater proportion of time when using a computer to record opinions than when using pen and paper ($p>0.05$ using Chi squared). There is no significant difference in self-assessment scores between Group C, who never used the DSGS v2 model, and Group D who used the DSGS v2 model for 18 weeks, when recording opinions of group skill usage on paper.

STAGE 2 SELF-ASSESSMENT	**PEN AND PAPER**				**COMPUTER**	
	GROUP C		**GROUP D**		**GROUP D**	
	Frequency	*Percent*	Frequency	*Percent*	Frequency	*Percent*
5 (all the time)	41	*43.62%*	61	*43.57%*	74	*59.68%*
4	30	*31.91%*	50	*35.71%*	32	*25.81%*
3	15	*15.96%*	17	*12.14%*	9	*7.26%*
2	7	*7.45%*	8	*5.71%*	2	*1.61%*
1 (none of the time)	1	*1.06%*	4	*2.86%*	7	*5.65%*

Table 6.4: Comparison of self-assessment scores

Over 80% of the children, regardless of the medium, gave a self-assessment in Stage 2 that was identical or only one mark different from their partner. This is shown in Table 6.5.

DIFF IN STAGE 2 ASSESSMENTS	PEN AND PAPER				COMPUTER	
	GROUP C		GROUP D		GROUP D	
	Frequency	*Percent*	Frequency	*Percent*	Frequency	*Percent*
0	27	*57.45%*	37	*52.86%*	39	*72.12%*
1	11	*23.40%*	26	*37.14%*	16	*19.39%*
2	9	*19.15%*	4	*5.71%*	3	*3.64%*
3	0	*0.00%*	2	*2.86%*	3	*3.64%*
4	0	*0.00%*	1	*1.43%*	1	*1.21%*

Table 6.5: Comparison of differences in self-assessment scores in Stage 2 of DSGS v2 model

Experience with the model has limited impact on the amount of differences in opinion. Using paper as the medium Group D are slightly more likely to differ in opinion compared to Group C ($p>0.1$ using Chi squared).

In Stage 3 children in Group D state on average they used group skills 67% of time when recording opinion on a computer, compared to 78% of the time when the medium is paper (see Table 6.6). This is a significant increase ($p>0.05$ using a t-test). There is no significant difference in overall assessment between Group C and D when the medium is paper. Significantly fewer children from Group D believe they follow group skills over 90% of the time when recording overall self-assessments on a computer compared to when using paper as a medium. These are both significantly less than the 47.4% of children in Group C who think they use group skills over 90% of the time ($p>0.01$ using Chi square).

AVERAGE ASSESSMENT FOR GROUP SKILL USAGE OVERALL	PEN AND PAPER		COMPUTER
	GROUP C (%)	GROUP D (%)	GROUP D (%)
Self-assessment in Stage 3 *Standard deviation*	82 *18.22*	78 *17.36*	67 *17.86*
Percentage of children whose Stage 3 self-assessment was greater than 90%	47.4	26.9	7.1

Table 6.6: Comparison of assessment in Stage 3, group skill usage overall

The low number of children in Group D believing they used group skills over 90% of the time when self-assessments were recorded on a computer contrasts with the Stage 2 assessments in Table 6.4. Almost 60% of the Stage 2 self-assessments show the children believed they used that group skill 100% of the assessment time. This lack of correlation between Stages 2 and 3, the self-assessment during the task and at task completion, occurs in all of the conditions and is illustrated in Figure 6.3.

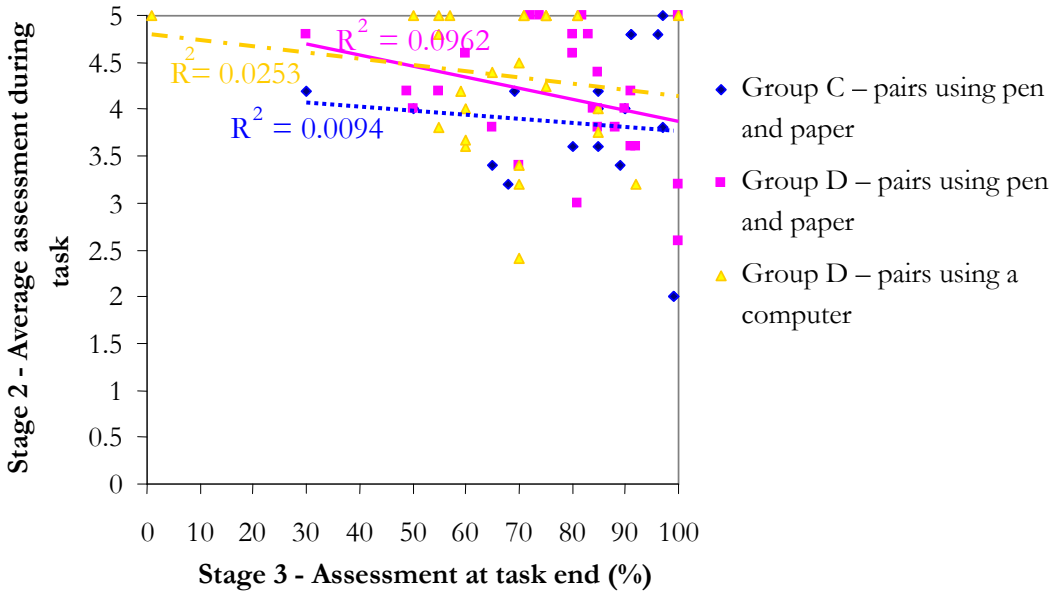


Figure 6.3: Graph showing the correlation between self-assessment during and at the end of the task in Study 4b

6.2.4 Discussion

As in Study 4a, children using a computer to record Stage 2 self-assessments are more likely to state they followed a rule all the time and give the same opinion as their partner compared to those using paper. This suggests the medium, or having a single display that allows observation of the other child's opinion, impacts the values recorded. Alternatively, it could be to do with the previous computer experience and believing the one "right" answer is required. The high number of perfect self-assessments in Stage 2 contrasts with the lower overall opinion in Stage 3, suggesting that recording opinions on a computer does not lead to a greater usage of group skills.

The lack of correlation between the Stage 2 and Stage 3 assessments implies the children are not more aware of behaviour throughout the task through using the DSGS v2 model. It could indicate the children may not have the metacognitive ability to relate previous actions to an overall assessment. This contradicts Study 4a. It may also relate to the environment and the hidden curriculum (Branco 2001). In this study, the children are using the model in class. In School X the emphasis of the teaching is on domain skills, in this case the TGT of programming robots. Studies 1, 2 and 3a show that the children focus on the TGT and minimal time is spent on the GST.

Repeated use of the DSGS v2 model has had some impact. Children used to reflecting on their behaviour, Group D, are more likely to disagree with their partner regardless of recording

medium. They also believe that they use group skills for a lower proportion of time overall, despite being more likely to state they follow a group skill perfectly during the task. It is unlikely that Group D are beginning to either internalise the terminology used by adults and relate it to this context, or accurately reflect on their own actions. If this were true, there would be a stronger correlation between their self-assessment during the task, and overall. Instead, it is hypothesised that the children are beginning to understand that self-assessment is subjective. This means they are able to apply their own criteria and understanding when self-assessing without having to first agree with their partner. The limited findings from these studies suggest that the training was insufficient to change the children's understanding of group skills. Adey and Shayer (1994, pp.60-61) argue that for an intervention program to have a permanent effect in changing the way that people think it need not be intensive, but that it must be maintained at a steady rate over a long period, they recommend 2 years.

6.3 Summary

The disadvantages of using a computer as the medium for the DSGS models given in Section 4.2.2 include the possibility that the children will not record their opinions honestly. This is due to previous ICT experience in which the children believe the correct answer or procedure is required. To ensure this preconception of computer use does not impact the self-assessments two studies were conducted.

The first was done outside of the classroom setting. The Stage 2 and 3 self-assessments recorded on paper were compared with those recorded on one computer. Children using paper are more likely to differ, give a wider range of scores, and spontaneously discuss differences between their opinions. Additionally, their assessments during the task correlate to their overall opinion of group skill usage. This suggests they are consistently, although not necessarily accurately, applying assessment criteria.

The later study was to provide evidence to support this finding. The results show that children are still more likely to differ in opinion and give lower marks in Stage 2 if using pen and paper. In Stage 3 those using a computer are more likely to give a lower overall assessment of the time spent following group skills. Unlike the previous study neither medium results in a consistency between assessments during the task and overall. This could be to do with the context. In the classroom the focus appears to be on the TGT instead of the GST.

Chapter 7

FEEDBACK

The aim of the software is to act as a training scheme that raises awareness of group skills. The group skills task (GST) is run concurrently with the teacher-given task (TGT). In both Developing and Supporting Group Skills (DSGS) models the children are required to agree criteria for assessing group skill usage (Stage 1) and prompted to justify opinions when discrepancies occur in self-assessment between group members during the task (Stage 2). However, these processes *do not* highlight misconceptions about group skill usage or definitions. As shown in Studies 1, 2 and 3 children overestimate group skill usage and have difficulties in applying criteria consistently. To assist the children develop an accurate awareness of behaviour and definitions of group skills feedback is given (Stage 4). This is in the form of:

- A numeric value generated by the model using the children's own self-assessments and time on the GST
- A prompt to focus discussion about any discrepancies between this and their own self-assessments

Feedback is a common method for assisting performance (Tharp 1993). It is used in existing training schemes, for example, when teachers give their opinions, or a discussion on group skill usage is held (e.g., Johnson *et al.* 1990, Bennett and Dunne 1992e). It is hypothesised that feedback will reduce misconceptions by: i) pointing out discrepancies between the self-assessments during and at the task end, and ii) contrasting their overall assessment with that recorded during the TGT. As children become aware of inaccuracies and discrepancies they will become better able to plan future interactions. This should be demonstrated by a reduction in discrepancies and more time spent using group skills. This understanding will be gained through appropriation; through talking and reflecting the children are expected to refine their understanding of group skills (Newman 1990, Vygotsky 1978).

This chapter describes how the feedback on group skill usage was generated in the Aliens implementation of the DSGS v2 model. Unlike the range of data discussed in Section 2.6 it looks at the sufficiency of self-monitoring as a means of accurately generating a numeric score representing group skill usage, and the impact this value and the prompts have on the children's behaviour.

7.1 Generation of the heuristic

The first goal of the feedback is to highlight inconsistencies in the self-assessment. Three types are possible: i) discrepancies *within* the group's self-assessments during the task (Stage 2), ii) discrepancies *between* self-assessments during the task and overall (between Stages 2 and 3), and iii) overestimation of group skills used. The second goal is that the feedback should accurately reflect the proportion of time spent using group skills. These two aims dictate what should be incorporated into the heuristic to generate the "score". Unlike existing CSCL systems that provide feedback on the amount of group skill usage (e.g., Soller 2001, Soller *et al.* 2002, Constantino-Gonzalez and Suthers 2001a, 2001b, Baker and Lund 1997) the DSGS v2 model does not have access to all the exchanges between the group members. Instead, this system uses the self-assessments of the children and the time taken between key presses.

The first aspect to consider is inconsistencies between the pair when assessing how much of the time a single rule was followed. If they are using the *same* criteria and considering the *same* period their opinions should be identical. The opinion generated should therefore take into account the actual self-assessment of both the children during the task and reduce in size as the differences in opinion increase.

The second area for inclusion is discrepancies caused by the children inconsistently applying criteria for measuring group skill usage between assessments in Stage 2 and in Stage 3. It is not sufficient for the heuristic to take into account just the assessments within Stage 2. The children could accurately report they never cheated, which would imply they used group skills throughout the TGT. Their Stage 3 overall self-assessment may be low however, as they realised they never listened to each other. In this example if the opinion was generated purely on Stage 2 assessments it would overestimate the group skills used and be unreliable. To prevent this type of error both assessments must be incorporated. If both scores, those during and at the end, are weighted equally, and the criteria for assessing group skill usage are applied consistently the two values will be similar. This causes the score generated to match the Stage 3 overall evaluation. Differences between the assessments will be reflected, but by weighting the impact on the generated opinion is reduced.

The third area to consider is overestimation of group skill usage. If the children stated they followed every rule all the time in Stage 2, then gave themselves 100% for group skill usage overall in Stage 3 there would be no discrepancies between or within their scores. Nevertheless, as shown in Studies 1a, 2 and 3b it is unlikely the children worked perfectly together and some reduction in the value generated is required.

The aim is for the opinion generated by the implementation of the DSGS model on the amount of group skill usage to correlate to the opinion of a teacher observing the children with access to other information, such as body language, conversation, and help requested. (The self-assessments are meant to reflect such information, but are reliant on the individual's understanding of group skills.) If large discrepancies exist between the children's opinions, they are required to establish why, and consequently refine their understanding. As the assessments become more consistent, the importance of the accuracy of the feedback increases. If the generated value states the children used group skills 99% of the time when they were working independently, for example, the children will be unaware they need to reconsider their interactions. Consistency within *and* between the scores will shift the emphasis of the feedback to prompting the children to plan how to improve group skill usage.

Using the self-assessments generated by Groups B, E and F in Studies 3a and 4a, an equation was developed through a process of inspection incorporating the requirements above. It correlates the information given by the children with the opinion of the teacher or researcher who assessed group skill usage. The correlation is most significant for those that used pen and paper to record their opinions, Group E ($r_s = 0.426$, $p>0.025$ using a Spearman test). In this case, the children were consistent between self-assessments during the activity and overall. Accurate feedback would assist on planning how much change is needed to work more effectively as a group. It is also significant for those who used a computer to record their opinions, Groups B and F ($r_s = 0.522$, $p>0.1$, $r_s = 0.509$, $p>0.1$ using a Spearman test respectively).

The first term in the equation relates to Stage 2, the second to Stage 3. Each contributes equally to the final score. The equation is given below:

$$\frac{\sum_{i=1}^{n}(5 - abs(C_1During_i - C_2During_i) * (C_1During_i + C_2During_i - \alpha)}{n} + \frac{(C_1Overall + C_2Overall - \beta)}{4}$$

where n Number of times Stage 2 assessments occurs (range 1-5)
C_1 Child 1
C_2 Child 2
$C_x During_i$ Child x i^{th} self-assessment during Stage 2 (range1-5) x = 1 or 2, i = 1 to n
$C_x Overall$ Child x overall self-assessment in Stage 3 (range0-100) x = 1 or 2

If $C_1During_i + C_2During_i = 10$ then $\alpha = 1$
Else $\alpha = 0.5$

If $C_1Overall + C_2Overall = 200$ then $\beta = 15$
If $C_1Overall + C_2Overall > 180$ then $\beta = 10$
If $C_1Overall + C_2Overall > 160$ then $\beta = 5$
If $C_1Overall + C_2Overall > 140$ then $\beta = 3$
Else $\beta = 0$

The constants, α and β, are required subtractions to compensate for the overestimation that occurs in self-assessment. They decrease proportionally with the self-assessments since a very high opinion of group skill usage suggests the children are overestimating, and if the opinion decreases it appears the children are reflecting more accurately on their behaviour.

The value the assessment during the task will be higher if *both* children's self-assessment is similar. This is because the model presumes they are using shared criteria. Suppose both children in the first assessment in Stage 2 state they followed the rule half the time and so chose 3. The value of $(C_1During_1 + C_2During_1 - \alpha)$ would be 5.5, and $(5 - abs(C_1During_1 - C_2During_1))$ would be 5 - resulting in a value of 27.5. If one child had chosen 5 – they had followed the rule all the time, and the other 1 – they had never followed the rule, the value of $(C_1During_1 + C_2During_1 - \alpha)$ would still be 5.5, but $(5 - abs(C_1During_1 - C_2During_1))$ would give 1 - resulting in a value of 5.5. This impacts the computer-generated score and should encourage the children to think about their self-assessments during the task to explain the differences.

The final mark generated by the Aliens software includes two values relating to the time taken. The first is a deduction if there is a small average time between a prompt for discussion and the key press to move to the next screen. It is assumed that the discussion necessary to develop an understanding has not occurred (Webb 1991). Marks are given for longer pauses, although a longer time does not imply appropriate discussion or reflection. The second value is based upon the time taken to produce an overall assessment in Stage 3. A short average time suggests they have not paused to reflect on the self-assessments in Stage 2 and marks are deducted. Again,

they are added if a long pause exists, although a long time does not imply reflection on the given task. These are shown in Table 7.1.

Average time Stage 2 discussion prompt displayed or average time taken to perform final self-assessment in Stage 3	< 15s	<30s	≥30s	>60s
Value	-5	-3	3	5

Table 7.1: Values added based on time prompts displayed

To illustrate how the software calculates a score Table 7.2 contains data logged over five self-assessments recorded in Stage 2, the overall assessment of group skill usage from Stage 3, and the values derived from them.

i =	STAGE 2 (DURING) 1	2	3	4	5	AVERAGE TIME PROMPT DISPLAYED	STAGE 3 (OVERALL)		AV. FINAL REFLE-CTION TIME
RECORDED VALUES									
$C_1During_i$	4	5	3	5	5	32s	$C_1Overall$	92	12s
$C_2During_i$	5	5	4	5	4		$C_2Overall$	87	
CALCULATED VALUES									
$C_1During_i + C_2During_i$	9	10	7	10	9		$C_1Overall + C_2Overall$	179	
α	0.5	1	0.5	1	0.5		β	5	
$abs(C_1During_i - C_2During_i)$	1	0	1	0	1				
$5 - abs(C_1During_i - C_2During_i)$	4	5	4	5	4				
$C_1During_i + C_2During_i - \alpha$	8.5	9	6.5	9	8.5				
ADDITIONAL MARKS									
						3			-5

Table 7.2: Self-assessment scores and calculated variables

Substituting the values from Table 7.2 into the equation with the time values gives:

$$\frac{(4*8.5)+(5*9)+(4*6.5)+(5*9)+(4*8.5)}{5}+\frac{179-5}{4}+3-5$$
$$=\frac{34+45+26+45+34}{5}+\frac{174}{4}+3-5$$
$$=\frac{184}{5}+\frac{174}{4}+3-5$$
$$=36.8+43.5+3-5$$
$$=81.3$$

The Aliens software would state the computer believed the children used group skills 81% of the time. However, this score is not sufficient as feedback in itself. As described in Section 5.1.4 this score is used with the children's own self-assessments to generate a comment and a question to encourage discussion or future planning.

Certain assumptions were made in developing the heuristic and about the impact that it will have on future behaviour. The first is that the children are capable of accurately monitoring their interactions. This is supported by: i) the development of metacognition allowing the children to reflect on their thinking (Sharples 1999, Azmitia 1998), ii) the existing use of questionnaires used to assess group skill usage in this age group (e.g., Bennett and Dunne 1992e), and iii) children being able to accurately apply assessment criteria (Finlayson and Cook 1998, Ager 1998, Study 3b). The second is that the children can perform both tasks simultaneously, i.e., they can complete the TGT, *and* spend time reflecting and discussing behaviour. The final assumption is that the children believe the computer-generated opinion. This is necessary for appropriation, i.e., understanding how their actions resulted in the computer-generated opinion.

7.2 Study 5: The accuracy of the equation and impact of the feedback

The remainder of the chapter considers: i) the correctness of the assumptions described above, ii) the accuracy of the heuristic itself, i.e., the correlation of the scores with the teacher's opinion, which strengthens as the consistency between assessments increases, and iii) if the feedback has had any impact on future behaviour. The results used for this are taken from Study 5 described next.

7.2.1 Participants and setting

Twenty-eight children, Group D, from School X participated in this study. The abilities of the pupils were similar to the National Average in mathematics, English and science using the QCA scores (see Appendix F for a comparison of the formal assessments for the schools). The study took place over 18 45-minute Information and Communication Technology (ICT) lessons in the school's computer suite.

7.2.2 Method

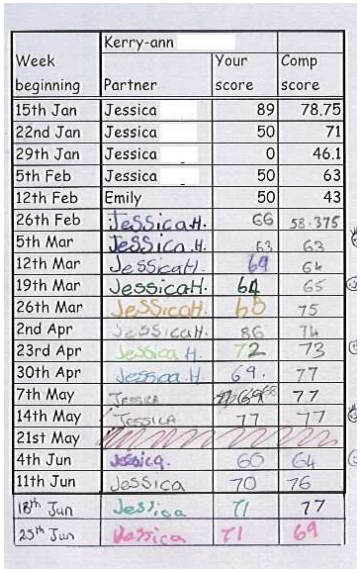
Each week the children used the Aliens software based on the refined DSGS v2 model (see Section 5.2). The TGTs used concurrently were designed to be joint activities–joint outcome tasks, but the applications did not enforce participation or test for individual understanding They were based on the Department for Education and Employment's Scheme of Work for ICT (DfES 2001). (Examples of lesson plans are given in Appendix J.) In self-selected pairs, they chose and agreed criteria for 3 rules for working together at the start of each session. During the lesson, the children independently self-assessed the time spent following one of their rules randomly selected by the computer. Based on these values, they were prompted for examples, reasons why self-assessments differed, or plans to improve their future behaviour. At the end of the lesson they were asked to assess overall usage of group skills. The resulting feedback was designed to encourage discussion to improve future group work.

Log files were generated for each child in each lesson. These contained:

- The rules and the time taken to choose each of them
- The rule assessed, the time taken, and score given for themselves and partner in Stage 2
- The time the discussion prompt screen was open
- If used, the time the message box prompting discussion was displayed
- The time taken and overall assessment for themselves and partner in Stage 3

In week 6 the children were given a record sheet to record their scores, see Figure 7.1. This was a table with four columns: the dates had been entered and they wrote their partner's name, their own opinion of how well they worked together, and the computer's value. This was done so the children could monitor their own progress and analyse how their behaviour affected the scores over time.

In informal discussions during the lessons the children were asked why the computer gave the marks it did and what they thought of the accuracy.



Week beginning	Kerry-ann Partner	Your score	Comp score
15th Jan	Jessica	89	78.75
22nd Jan	Jessica	50	71
29th Jan	Jessica	0	46.1
5th Feb	Jessica	50	63
12th Feb	Emily	50	43
26th Feb	JessicaH.	66	58.375
5th Mar	JessicaH.	63	63
12th Mar	JessicaH.	69	64
19th Mar	JessicaH.	64	65
26th Mar	JessicaH.	66	75
2nd Apr	JessicaH.	86	74
23rd Apr	JessicaH.	72	73
30th Apr	JessicaH.	69	77
7th May	Jessica	69	77
14th May	Jessica	77	77
21st May			
4th Jun	Jessica	60	64
11th Jun	Jessica	70	76
18th Jun	Jessica	71	77
25th Jun	Jessica	71	69

Figure 7.1: Record sheet

One pair in every lesson was tape-recorded. During six of the 18 lessons the teacher assessed the proportion of time spent using group skills. These were then compared with the marks provided by the DSGS v2 software heuristic described in Section 7.1. The results of this are given in Section 7.4.

At the end of the study, the teachers were asked to complete a questionnaire about the group skills displayed by the pupils in their class. The results from the children being given the feedback, and the impact the feedback had on their behaviour are given in Section 7.5.

Caution is needed when drawing conclusion from this data since only two teachers and 28 children participated and a restricted number of groups were tape-recorded. Moreover, the opinions of the teachers are subjective, they could have used their knowledge of the child, previous experience etc., when assessing group skills displayed during the ICT lesson. Finally, the logged information need not necessarily represent the time "on task".

7.3 Testing of assumptions

There were three assumptions made when the heuristic was generated – that the children: i) were accurately monitoring their behaviour, ii) were performing the two tasks concurrently, and iii) believed the computer-generated feedback. This section uses the log files, observations and recordings to assess the accuracy of these statements.

7.3.1 Results

The observations and recordings suggest that many children were *not* monitoring their behaviour as requested. The children appeared to focus on the TGT, and minimal time was given to the GST. An example is the addition of rules for working well together. Out of the 693 rules selected by the pairs, only 31 rules were added. Of these, five were copies of existing rules, although only four added rules were spurious, e.g., "work, work, work". In the recordings and observations, children did ask each other before choosing rules but rarely discussed criteria for the assessment. The log files corroborate this. The average time the screen with the rules was open was seven seconds. This time is insufficient for a discussion of assessment criteria.

When performing the assessments in Stage 2, many children considered behaviour in general instead of one group skill. Those that did this were likely to become frustrated at the interruptions: "but Miss…we've done that". The average self-assessment during the task in Stage 2 is 4 (standard deviation 0.4), i.e., the majority of children believed they followed rules for group work three-quarters of the time. The children gave different self-assessments on the amount of time a rule was followed in 21% of recorded opinions. Over the 18 weeks, the overall assessment for group skill usage averaged 66.7% (standard deviation 7.0). The computer generated score averaged 61.3% (standard deviation 5.5). On average 16.7% of the children assessed themselves as using group skills more than 90% of the time. This fell to 4.4% of the children in the last four weeks.

The recordings suggest the prompts did not generate much discussion. If they occurred, they appeared to be statements of fact rather than suggestions, as shown in the extract below:

Pupil 1: You put three?
Pupil 2: No… I put one
Pupil 1: Why?
Pupil 2: 'Cos we didn't swap over
Pupil 1: But…
Pupil 2: We didn't.

This is reflected in the time the Aliens software was displayed. In the first week, the screens were viewed for an average of 335 seconds. In the last, 197 seconds. There was a reduction in the times taken to choose rules, and self-assess in Stages 2 and 3. The length of time for discussions in Stage 2 remained similar due to the message box prompt reminding them to talk; 114 seconds in the last week compared to 125 seconds in the first. The message boxes appeared effective in encouraging discussion. The children were concerned that the computer had actually been listening to them, and were observed discussing the group skills as requested. This is reflected in the time this message is displayed. Despite the option to close it immediately these prompts were open for an average of 22.9 seconds, normal discussion prompts were displayed for average of 14.3 seconds.

The prompts were also frequently disregarded in Stage 4. Instead of considering methods to improve future performance, the children would compare scores. Over the study, the area of competitiveness shifted from getting a high score to entering an opinion in Stage 3 similar to the

value returned; the phrase "I got it right!" was frequently recorded in later weeks. The record sheets did *not* perform the intended role as a basis for discussion and reflection enabling the children to analyse how their behaviour would cause a change in the scores.

When asked, the majority of children felt the computer's opinion was more accurate than their own, with many believing the computer had been monitoring them. Despite the cameras used for video conferencing being off, some children believed they controlled the generation of the message box reminding them to talk. One pair of pupils reasoned their low score was due to the fact the computers were networked. It had realised they could not have worked together as one boy had logged onto another machine. Only one pair was recorded actively disagreeing with the computer. They told the researcher they had worked together and concluded the computer had a virus.

7.3.2 Discussion

The results suggest that the children can monitor behaviour independently using their own criteria since they are willing to give different opinions to their partner in around 21% of the self-assessments. This suggests they are taking responsibility for their opinions. An improvement in accuracy is also supported by the reduction in children who give high overall assessments. In Study 2, 39.1% believed they used group skills over 90% of the time. In Study 4a it was 50%. In these sessions, the average was 16.7% and was falling with use of the DSGS v2 model. There are insufficient recordings and observations to confirm that the self-opinions of the children reflect what occurred though.

Performing the two tasks concurrently still appears problematic. The children spend little time on the GST activities. The maximum time Aliens was on screen was only 335 seconds in 45 minutes. Some reduction in display time over the study could be attributed to the children being more aware of what was required, not reading the instructions, referring to previous examples, and using existing agreed criteria. Alternatively, the child with the mouse may be acting as the dominant group member and unilaterally deciding to continue without consulting their partner. Regardless of this, to spend just over three minutes in discussion or reflection of rules, criteria, individual group usage and overall demonstration of group skills in the last week suggests the children's main focus is the TGT.

This is supported by the lack of discussion on group skills recorded. One explanation is the context. Within the classroom there is a focus on domain knowledge rather than general learning skills (Galton 1998, Meittinen 1999). The short reflection and discussion times on group skills reflect the classroom observations in Study 1d. In these the children initiated less than 4% of discussion, and less than 20% of the time was spent responding to questions. The encouragement to talk about group skills, and explicitly plan contrasts to the perceived desired behaviour of silence and closed questioning. The lack of effort on the GST would also explain why the prompts failed to generate the talk noticed in other research (e.g., Enyedy *et al.* 1997, Wegerif 1996b, Pilkington 1998). These prompts focused on the main activity, the TGT. When the assessment activity was the TGT in Study 3b children could reflect on group skills using the prompts as a basis for discussion. This implies that the children cannot allocate equal effort to both tasks and concentrate on the domain knowledge, in this case ICT as a curriculum subject. It appears children do not view computers as mediating artefacts to promote a discussion and an understanding of group skills.

This focus on the TGT would explain the lack of discussion in the group relating to the feedback given. It would also explain why children do not try and appropriate the feedback through actively reflecting and reasoning about how the score was generated. They are not able to create the context in which the information can be appropriated (Newman *et al.* 1989) despite the perceived expert and reliable status of the computer. This lack of reflection and discussion of group skill usage could also be due to a lack of metacognitive awareness. The children may not be able to perform this reasoning in this context, or they may have difficulties with the switch to reflecting on their behaviour and providing justifications for their actions. The representation may add to this problem. The existing perception of computer scores is they are a summative value relating directly to the knowledge entered into the computer. Using a numeric representation for the opinion of the time spent using group skills fulfils the criteria of being an easily recognisable representation (Gallimore and Tharp 1990); but appears to be interpreted as an accurate record of behaviour instead of focus for discussion.

7.4 The accuracy of the equation

In order for the appropriation process to be meaningful the "score" has to be reliable. If the system generates a high figure when the children worked independently the child may not infer group skills were absent. As described in Section 7.2 the formula was incorporated into the Aliens software and used in 18 ICT lessons. The teacher of the class assessed the group skills displayed by the pairs for six of these lessons using the criteria described in Study 2, i.e., body language, conversation fragments, help requested etc. This section investigates whether the scores given by the teacher correlate to those generated by the software based on the children's self-assessments of this data.

7.4.1 Results

The heuristic was designed to highlight discrepancies, and reflect more accurately the teacher's assessment on group skill usage if the Stage 2 and 3 self-assessments are more consistent. To test this an analysis was made based on the consistency between the self-assessments and the opinions of the teacher. For the six weeks that the teacher had assessed the proportion of time the pairs had spent using group skills the children were divided into two groups. Those considered *consistent* had a difference of less than 20 marks between their Stage 3 assessment, and their Stage 2 assessments summed and represented on a scale of 0 to 100. The *inconsistent* group had a difference of more than 20 marks between these figures. These were then correlated with the opinion of the teacher and the results are shown in Table 7.3.

	WEEK					
PEARSON PRODUCT	**6**	**11**	**14**	**16**	**17**	**18**
Teacher's assessment correlated to computer generated score for inconsistent students	0.6367	0.2690	0.6353	0.5295	0.4590	-0.0484
Significance	*p>0.05*	-	*p>0.01*	*p>0.05*	*p>0.1*	-
Number of pupils	*12*	*13*	*17*	*16*	*16*	*16*
Teacher's assessment correlated to computer generated score for consistent students	0.6353	0.4567	0.4592	0.6821	0.7329	0.4302
Significance	*p>0.02*	-	-	*p>0.05*	*p>0.02*	-
Number of pupils	*16*	*11*	*11*	*12*	*12*	*12*

Table 7.3: Correlation of consistent self-assessments with the opinion of the teacher

The table shows that the score generated by the computer is significantly similar to the opinion of the teacher half the time. Apart from week 14, there is a trend for a stronger correlation

between the teacher's opinion and the consistent children. The relationship is not exact. In weeks 11 and 18 there are no significant correlations between either group and the teacher.

7.4.2 Discussion

Table 7.3 indicates the approach of using self-assessments to generate feedback is promising. A significant relationship exists between the teacher and computer generated score in over half of the sessions, despite Section 7.3 highlighting the minimal time spent on self-assessment. Furthermore, in half the recorded sessions the formula is more accurate for children who appear to be applying the criteria for assessing group skill usage consistently. This suggests they are using an agreed set of criteria, which could be similar to the definitions to an adult (Wertsch and Stone 1985).

The formula needs refining. The marks based on time, for example, should be changed as the short discussion times (these are given in Appendix K which contains a summary of the log files over the course of the 18 weeks in a tabular form showing the averages) suggest units of 15 seconds are too great. It also requires further testing before any conclusions can be drawn. As stated, it is based on a small sample size with subjective data. To ensure the accuracy of the teacher's opinion, for example, several teachers should assess the use of group skills by the children.

7.5 The impact of the feedback

If the feedback is having an impact on the understanding of group skills, the overall self-assessment and feedback from the software, which is based on the self-assessments in Stages 2 and 3, should correlate. This is because the children develop shared criteria for assessing group skill usage and an awareness of their behaviour. It is hypothesised an improved correlation would be reflected in an improvement in the group skill usage. However, the children tend to overestimate the time spent using group skills, so improvements in behaviour may not be reflected in their self-assessments. In fact, these self-assessments may *decrease* as the children refine their understanding of group skills and conclude they are not following a rule as well as initially believed. Instead, to test this claim the opinion of the teacher about the child's ability to work in a group is compared to this correlation to determine if a relationship exists.

7.5.1 Results

The correlations between the children's overall opinion and the feedback score generated were plotted for the first five weeks and the last five weeks. Seventeen out of the 28 became more consistent in self-assessments. This is shown by a stronger correlation between their own final opinion and the computer feedback in the last five weeks than the first five. Out of the 17 children who had improved group skill usage according to the teachers' questionnaire 13 had become more consistent. An example is Lyndsey. According to her teacher, her group skills improved the maximum amount possible. Figure 7.2a shows that initially there was a weak correlation between the computer-generated mark and Lyndsey's own overall opinion. As the study progressed she became more consistent in her assessment during the task and at the task end, resulting in the improved correlation with the computer score.

Out of the 11 children that stayed the same in group skill usage according to the teachers' questionnaire, nine displayed less correlation between their own overall self-assessment of group skills used and the value generated by the computer. Daisy, who had an average QCA score of 45.42 compared to Lyndsey's 45.25, showed no improvement in her ability to work in a group according to her teacher. Figure 7.2b shows that in the first five weeks there was a high correlation between her overall assessment and the computer opinion of the amount of group skills used. In the last five weeks her self-assessments in Stages 2 and 3 were inconsistent, so her overall opinion no longer correlates to the computer score.

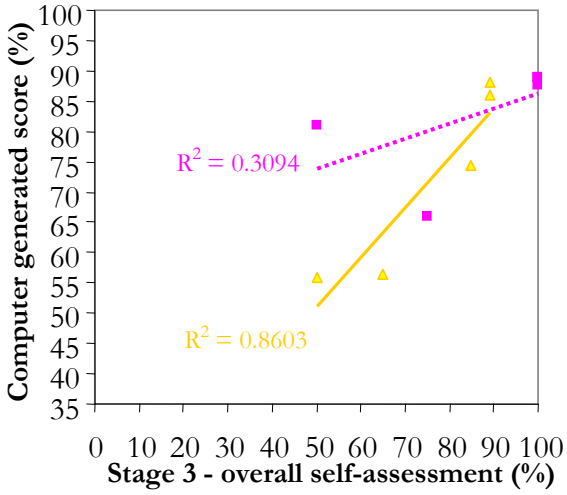
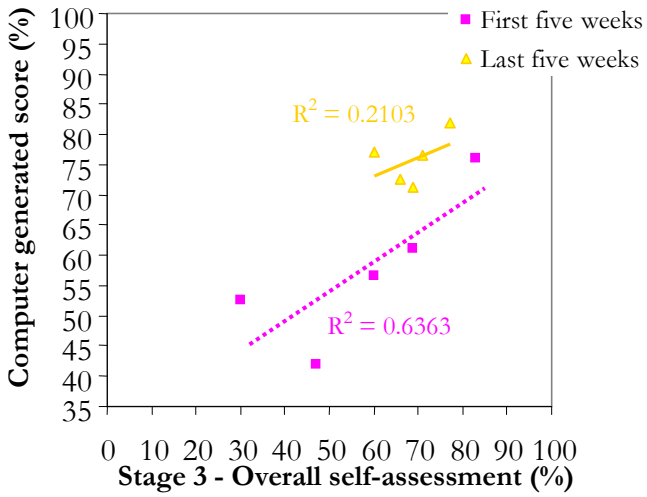


Figure 7.2a: Lyndsey's self-assessment



b: Daisy's self-assessment

This pattern of improvement or staying the same in group skill usage in the opinion of the teachers corresponding to an improvement or decrease in the correlation of the self-assessments during and overall with in the child is followed by 77.8% of the children.

The improvement in consistency does not relate to an improvement in the overall self-assessment of group skill usage. Eight out of the 13 children who improved in the eyes of the teacher recorded a drop in overall self-assessments between the first and last sessions, from 79% (standard deviation 17.9) to 60% (standard deviation 17.4). While seven out of the nine who recorded a greater discrepancy between the assessments during the task and at the task end, logged an improvement in overall self-assessments, from 50% (standard deviation 10.4) to 71% (standard deviation 14.5).

7.5.2 Discussion

Despite the lack of discussion shown by the short times in Section 7.3.1, many children appear to be influenced by the qualitative marks generated. The record of assessment may have some influence, as the children could base the current week's mark on the computer-generated mark from the previous week. This does not explain the increased consistency between assessments during the task and overall, and the relationship with the teacher's opinions. Over time, those that improved in the teacher's opinion had an improved correlation between their overall self-assessments and the score generated by the system. Conversely, in the majority that stayed the same the correlation decreased. This suggests that awareness of group skills occurs outside the ICT lesson but is reflected in self-assessment. Those children that become more aware of group skills are better able to assess their usage, which leads to a consistency in the application of criteria for judging the proportion of time they are used. If this improvement is based on reflection, or appropriation, then the findings suggest these processes are internal. This is concluded from the scarce discussions (Section 7.3), which suggests communication between children is less responsible for the change.

7.6 Summary

It was hypothesised that feedback could be generated from the self-monitoring performed by the children. For this to reflect group skill usage it had to take into account: i) the discrepancies *within* the group's assessment during the activity, ii) the discrepancies *between* the overall assessment and that during the activity, and iii) overestimation of group skills. Using this data a score for overall group skill usage can be calculated that correlates to the opinion of the teacher. As discrepancies between opinions decrease the score correlates more strongly to the opinion of the teacher to assist future planning.

The opinions recorded do not reflect accurate self-monitoring. They suggest that the children are not able to perform the TGT and the GST simultaneously – more effort is spent processing the TGT. It is hypothesised this is due to the children's expectations. Given their experience in the classroom, and the perception of the computer as a tool rather than an artefact to mediate group work, the goal is seen as the completion of the TGT. Their focus with respect to the GST is generating an opinion that is similar to the feedback "score". The children believed this value on overall group skill usage was accurate. This belief that the computer is an expert may assist appropriation. However, the misconceptions about how this information is obtained suggest they have insufficient cognitive ability to internalise these reasons and so amend future behaviour.

The failure of the prompts to promote discussion may also be related to the difficulty in performing two tasks. In existing research, they are beneficial as they refer to the TGT; in this activity, they relate to the GST, which is seen as less important.

The feedback does not necessarily reduce the inconsistencies in the Stage 2 and 3 self-assessments. This contradicts the hypothesis made within the model. A comparison with the teacher's opinion on ability at group work suggests children aged 9 and 10 have difficulty in recalling or assessing their own behaviour at the end of an activity – despite prompts to consider their behaviour throughout. Although the feedback does not appear to influence group skill usage, the software *does* reflect an awareness of group skills. The majority of children whose ability at working in groups improves according to their teacher show an improved correlation between overall self-assessments and the feedback based on opinions recorded during the task and overall. While the majority of children that stayed the same do not display this relationship.

Chapter 8

THE INTERVENTION PROGRAMME

Study 1 had shown that one session focusing on group skills was insufficient to affect group behaviour in the classroom. Other research shows behaviour is changed through a long exposure to a training scheme (Adey and Shayer 1994, Dawes *et al.* 2000). Study 2 suggested cognitive understanding is benefited by encouraging children to focus on group skills through a combination of reflection, responsibility and communication whilst performing the teacher-given task (TGT). Study 3a showed that enforcing self-assessment thus making the focus on group skills compulsory did not impact upon the TGT performance. These results suggest the Developing and Supporting Group Skills (DSGS) v2 model would benefit the task performance *and* raise awareness of group skills if used repeatedly. To investigate this suggestion an intervention programme was conducted over a period of 18 weeks. It also addressed the second issue of whether raising awareness of group skills led to them being formalised and transferred to other TGTs where the Group Skills Task (GST) was not explicit (GST) (Kozulin 1986, Vygotsky 1987).

This chapter describes the intervention programme, Study 6. It was hypothesised that learning gains would be made individually and as a group by using the DSGS v2 model. These can be cognitive, measured by task performance; and social, measured by use of language. Both types of information are required, as neither completely describes the outcome of the training scheme (Howe 1991). The results would be corroborated by the teachers' opinion.

The four hypotheses that relate to the expected improvement, based on existing studies and previous research, are summarised in Table 8.1.

	HYPOTHESIS	PROPOSED SUPPORTING EVIDENCE
a	The children who had used the DSGS v2 model would retain more of the information they had studied in the ICT lessons (as in Study 2).	Higher scores in an individual ICT quiz covering material learnt in the course of the study.
b	Due to a greater awareness of group skill usage, children who had used the model would work better as a pair or in a group of six compared to the control groups.	The group puzzles would be completed in fewer moves, with less help and errors as the children took responsibility for the performance and reflected before moving. Improved communication and reflection would be indicated by longer pauses between moves.

	HYPOTHESIS	PROPOSED SUPPORTING EVIDENCE
c	From internalising the prompts and practising reflection, responsibility and communication, those children who had used the model would demonstrate an increase in exploratory talk at the end of the study compared to the control groups.	Transcripts of the sessions would show more justifications, longer statements, less interruptions and less simultaneous speaking.
d	Teachers would identify an improvement in children who used the DSGS v2 model's general ability to work together within the classroom.	Questionnaire completed by the teachers at the end of the study.

Table 8.1: Four hypotheses to be investigated

The remainder of this chapter describes the study format and puzzles used, followed by an evaluation of the four hypotheses.

8.1 Study 6: Format

As shown in Figure 8.1 two classes were divided into matched groups to form two sets of equal ability. The children then participated in the three phases in Study 6:

- Phase 1 – pre-test in the form of puzzles (the children solved two out of four puzzles - one in a pair and one as a group of six)
- Phase 2 – system use, an introductory session was given on the importance of group skills, appropriate rules and criteria for assessment, then an identical set of ICT lessons with one set of children using the Aliens software to support the GST, and the other acting as a control
- Phase 3a – post-test in the form of puzzles (in the same groups of two and six the children solved the remaining two puzzles)

  Phase 3b – an individual quiz for testing recall of information studied

The groups performing the puzzles in Phase 1 and 3a consisted of children where some, none, or all would or had used the Aliens software in Phase 2. At the end of the study, the teachers completed a questionnaire about the level of, and any changes within, the group skills displayed by the individual children. The quiz performance, the group task performance, the recordings of exchanges while performing the group tasks, and the teacher questionnaires were used to investigate the four hypotheses. A triangulation process was built into the data gathered. An improvement in task performance, for example, should correlate to an improvement in the amount of exploratory talk occurring.

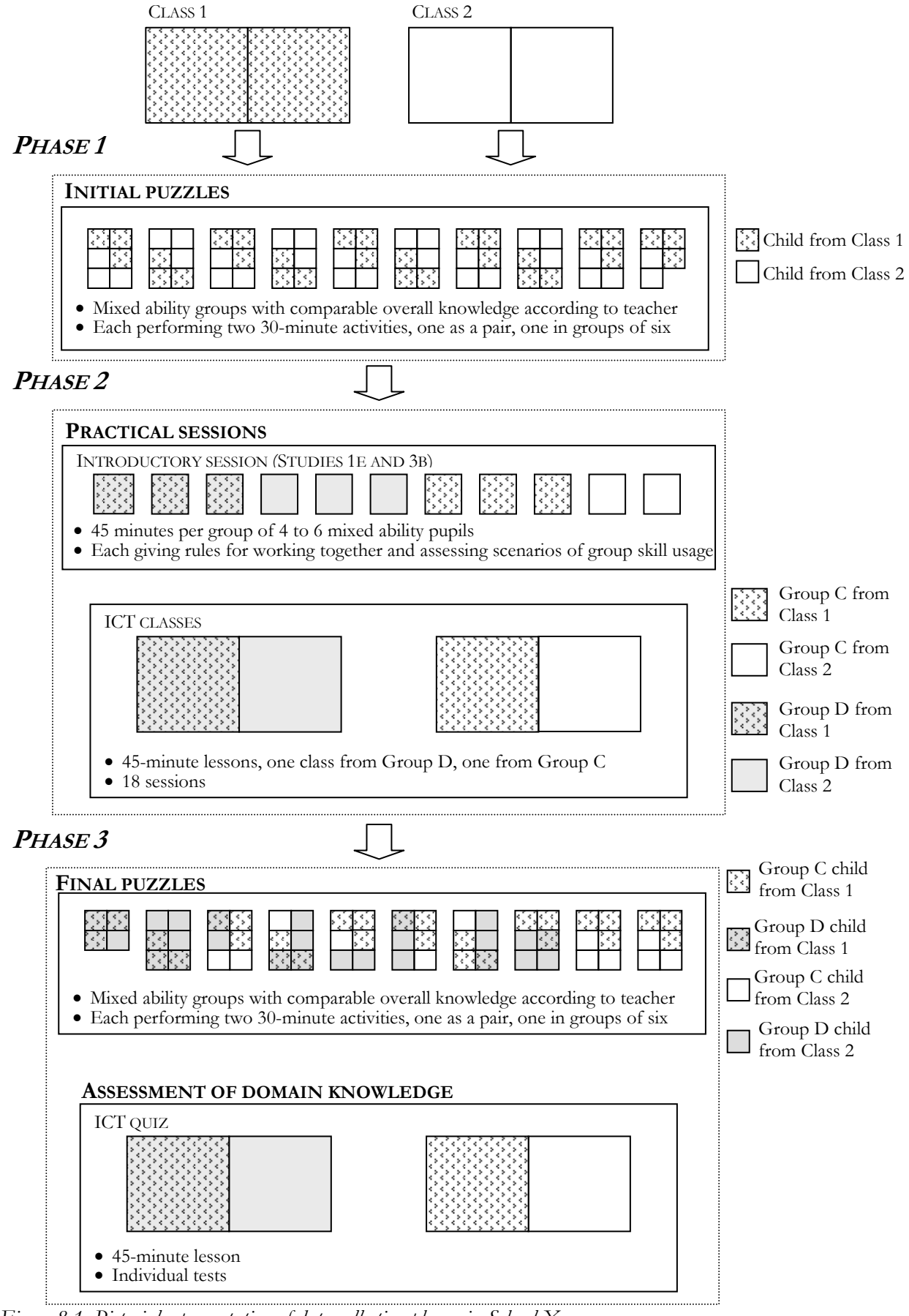


Figure 8.1: Pictorial representation of data collection phases in School X

Initially the study was to be run in Schools X and Y. Both schools had participated in Studies 1e and 3b. In the Easter term the computer suite in School Y was unavailable for six weeks due to refurbishment. Thereafter access was restricted, requiring the lessons to be given to groups of six instead of as a class. School Y also had pupil attendance difficulties. In Phase 1 48 children participated, in Phase 3a only 42. Absences meant only eight pairs were the same in the first and final set of puzzles, and only two groups of six. Consequently, the results given below are taken only from children in School X, although similar trends were recorded in School Y.

8.1.1 Puzzle summary

This section describes the four puzzles that provided quantitative and qualitative group performance information. The eight 9 and 10 year olds who had assisted in the design of the Aliens software also aided the puzzle development. They consisted of two girl-girl pairs, one boy-boy pair and one mixed pair. The goals were to: ensure the level of difficulty was appropriate, the wording could be understood, and it was possible to record the number of moves (Druin 2002). The tasks were termed puzzles, as the words "computer games" may suggest the tasks as more appropriate for boys, and cause the girls to disengage (Light and Littleton 1999d). They were presented as joint activity–joint outcome tasks, but there was no control over how the group interacted when solving them.

Two puzzles were designed for groups of six. The Statue is isomorphic to the Tower of Hanoi problem. It is equivalent to six rings and four rods. The children are told they are archaeologists in Peru and have to move a statue from its site to their camp. The statue had been broken into six pieces. The head is the lightest piece, and they became heavier towards the feet. Only one piece can be moved at a time, and a heavier piece can never be placed on a lighter one or it would break. The ground between where the statue was found and the camp is not solid. Only four places exist, including the camp and starting point, which are able to support the statue's weight. The Frogs puzzle is of equal complexity. It begins with six alphabetically labelled frogs, each on a lily pad, sitting in a row. Initially there are frogs A, B, C, an empty lily pad, then frogs D, E, and F. The goal is to exchange their places so that the order is D, E, F, an empty lily pad, A, B and C. Frogs A, B and C can only move right. Frogs D, E and F can only move left. Frogs can only move to an adjacent lily pad if it is free, or jump over one adjacent frog if the lily pad on the other side is free. Both puzzles for six are designed to take half an hour, they have a physical set of props, and the written instructions suggest the children could pretend to be a frog or statue

piece in order to plan their moves. (The instructions and prop descriptions are given in Appendices L1 and L2.) The task performance, measured by the number of moves, restarts and assistance required, was taken from observations.

These puzzles are far transfer tasks. They are different in domain, environment, and group size to the training scheme situation, i.e., they are not ICT activities, they are not computer-based, and the group is six not two.

The remaining puzzles were designed for pairs and should take under half an hour. The first is isomorphic to the Missionaries and Cannibals problem. In the Return to Earth puzzle three space policemen have to escort three alien criminals from the planet Zog to their spaceship using a apace shuttle. There are three rules: (1) the number of aliens can never outnumber the police either on Zog or the spaceship, (2) the shuttle needs somebody – either a policeman or an alien to fly it – and (3) the shuttle at most can hold two characters. A clue is offered at 6 moves, 12 moves and 18 moves. The second is the Petshop puzzle. This is isomorphic to the Tower of Hanoi task. The children are instructed to move four fish from a small tank to a large one. They are given two rules: (1) they can never put a large fish in a tank with a small one already present, and (2) they can only move the smallest fish in the bowl. If they succeed, they are given the option to do the same task with three bowls. The first exercise is equivalent to four rods and four rings; the continuation exercise is equivalent to three rods and four rings. (A full description of the puzzles and screen shots can be found in Appendices L3 and L4.)

To overcome the problem of self-recording found in the design phase these puzzles were implemented on a computer. If rule (1) in the Petshop puzzle or Return to Earth was violated, the rule was reiterated and the children had to restart the task. If the children broke one of the remaining rules, the system explained why this was illegal, and did not make the move. The reiteration of the rules provided cognitive offloading, allowing the children to focus solely on solving the task and working together (Kutnick 1994). The computer implementation logged the moves, illegal moves, times, restarts, and assistance given.

These puzzles were near transfer tasks because although not a curriculum defined Information and Communication Technology (ICT) task the children were working at a computer in pairs.

The children who participated in the design stages failed to note the similarity between the Petshop puzzle and The Statue, even when they were performed consecutively. This supports Pierce and Gholson (1994), who found junior school children were influenced by the physical appearance of the task.

8.1.2 Participants and setting

Two Year 5 classes of 30 pupils from School X participated in Study 6. The children had been observed as part of Study 1d. Based on the Qualifications and Curriculum Authority (QCA) Year 5 examinations the children's average ability was similar to the National Average for mathematics, science and reading, and they were above average for spelling (see Appendix F for a comparison of the formal assessments for the schools).

Before the study commenced each teacher ranked the pupils in their class by ability. Then taking the odd numbers from one class and the even numbers from the other, two matched sets were formed. This minimised the impact of other variables, such as teaching styles. One half formed Group C, the control set. The other, Group D, used the DSGS v2 model implemented as Aliens. Four pupils left in the spring term and two joined, resulting in Group C having 30 pupils and Group D having 28. Due to this, and inaccurate initial grading when compared to the QCA scores, the groups were not significantly different but Group D was more able than Group C. This is shown in Table 8.2.

	AVERAGE ABILITY (%)		
	Mathematics	**English**	**Overall mean**
Group C	46.99	45.43	46.06
Group D	50.89	44.03	47.46
T-test	*0.274*	*0.886*	*0.472*

Table 8.2: Comparison of groups by ability

8.1.3 Method

As shown in Figure 8.1 on page 157 and elaborated below Study 6 was divided into 3 phases: Initial puzzles, Practical sessions, and Final puzzles and Assessment of domain knowledge.

PHASE 1: INITIAL PUZZLES

Due to an absence, 59 children took part in Phase 1. They were divided into nine groups of six and one of five, with a similar overall ability according to the teacher's opinion. Groups of six

were subdivided into three pairs with an overall equal combined ability. They performed the puzzles at the end of the winter term. Each child performed two half hour activities in the library: one as a pair, one as a group of six.

The groups of six were arranged so that five did the Frogs puzzle, and five the Statue. Of those five groups given Frogs, three groups did the Return to Earth puzzle, and the remaining pairs the Petshop puzzle. Three groups did the Statue and the Petshop puzzle, and the remaining groups the Statue and Return to Earth. Half the groups started with a puzzle for pairs first. This is summarised in Table 8.3.

GROUP	1	2	3	4	5	6	7	8	9	10
Puzzle 1	Frogs	Return to Earth	Frogs	Petshop puzzle	Frogs	Return to Earth	The Statue	Petshop puzzle	The Statue	Petshop puzzle
Puzzle 2	Return to Earth	Frogs	Return to Earth	Frogs	Petshop puzzle	The Statue	Return to Earth	The Statue	Petshop puzzle	The Statue

Table 8.3: Initial group puzzle format

In all activities, the number of moves and amount of help requested were logged. All the puzzles performed in groups of six and half the pair activities were videoed. To minimise the Hawthorne effect the cameras were present in all cases so the children are unaware of which sessions were recorded.

PHASE 2: PRACTICAL SESSIONS

This was divided into two stages: an introductory session, followed by the ICT lessons. The introductory session was a discussion of the importance of group skills, appropriate rules and criteria for assessing whether these were followed, and scenarios to assess for practice. These activities are described in Studies 1e and 3b. For children in Group D this introduction was important. Study 3a had shown that children would ignore the introductory screens without support. These are essential screens, as interactions are enhanced if the children understand the rationale (Lin 2001, Kramarski and Mevarech 1997, Salomon 1984). The session gave children in Group D the opportunity to practise using the Aliens software with the researcher available to assist and answer any questions. It enabled the researcher to clarify any misunderstandings in the prompts and ensure they were not ignored. The session tried to scaffold the children in developing criteria for assessing how well rules were followed. If they chose "listen to each other" for example, they were prompted to explicitly state how they would recognise if this rule

was followed. Their responses included statements like: "don't interrupt", and "never speak at the same time".

Group D used the Aliens software based on the DSGS v2 model described in Chapter 5. It is domain independent, and was used in conjunction with the suggested work plan for ICT (DfEE 1998). ICT was chosen as: i) it was the only subject in which the children had regular access to the computers; and ii) the schools agreed as it fulfilled their curriculum requirements. Identical ICT lesson plans were used for structuring all computer sessions for both sets for 18 weeks between January and June. Each lesson lasted 45-50 minutes and was given by the researcher to reduce the impact of teaching styles. A complete list of the subjects covered and example lesson plans are found in Appendices M and J. As far as possible, the lessons were integrated into activities outside the classroom. In Unit 5F, environmental monitoring, the children observed the weather, the cloud cover, wind direction, and type of cloud outside the ICT lessons. This real data was used to create and test their hypotheses about the relationship between these factors. This was to make a real, or an authentic, task (Enyedy *et al.* 1997). These tasks were designed to encourage collaborative or collective working, that is, be joint activity–joint outcome tasks, but within the lesson there was no individual testing or requirement that both should participate.

As the model is reliant on encouraging discussion to raise awareness of group skill usage, the children were allowed to select their own partners. This contrasts with previous studies on collaboration, where it is suggested the most learning occurs when the participants differ by one degree of ability. Webb, for example, found most improvement when a pair consisted of a highly able child and one of middle ability, or a middle ability and low ability child (Webb 1985). This choice was given to the children as factors other than ability, for example, friendship, can dictate the amount of talk (Puntambekar 1995, Issroff 1999).

PHASE 3A: FINAL PUZZLES

Due to absences and children leaving and joining the classes, only 24 of the 29 pairs and six of the ten larger groups were identical between Phase 1 and 3a. However, they were matched in that the appropriate number had used the DSGS v2 model and the overall ability of the group was similar.

The children performed the two puzzles not yet attempted in Phase 3a. They worked in the library. If they had begun with the puzzle for pairs in Phase 1, they began with the task for six and vice versa. The actual puzzles and ordering are shown in Table 8.4. The groups were structured so that none, two, three, four or all had used Aliens, as shown in Figure 8.1.

GROUP	1	2	3	4	5	6	7	8	9	10
Puzzle 1	Petshop puzzle	The Statue	Petshop puzzle	The Statue	Return to Earth	Frogs	Petshop puzzle	Frogs	Return to Earth	Frogs
Puzzle 2	The Statue	Petshop puzzle	The Statue	Return to Earth	The Statue	Petshop puzzle	Frogs	Return to Earth	Frogs	Return to Earth

Table 8.4: Final group puzzle format

In all the activities the number of moves and amount of help requested was logged. Video cameras were present for all activities, but only the same groups were recorded.

PHASE 3B: ASSESSMENT OF DOMAIN KNOWLEDGE (THE ICT QUIZ)

The children were asked to individually complete a computer based quiz covering the ICT material taught throughout the terms. The quiz offered dynamic assistance. This was developed with two children who completed the test working with the researcher. On the basis of the help requested and given, prompts were built into the questions (Brown and Campione 1987). Initially the response to a mistake was a request to "try again and think carefully". These became more detailed with further errors, finally becoming structured suggestions listing the steps to be taken. If the child could not answer the question with this assistance, another question was presented. (The questions, and the assistance built into them, can be found in Appendix N.)

The quiz contained 17 questions, and the children were given up to 45 minutes in which to complete it. The questions were given in a random order to reduce copying. Due to absences and involvement in the quiz design phase, only 51 children participated.

The teachers completed a questionnaire at the end of Study 6. This contained their impressions of how the pupils in their class initially worked in groups and whether their performance had changed over the course of the study.

8.1.4 Limitations of the study

The advantage of integrating the study within a classroom environment instead of a controlled laboratory setting ensures the findings incorporate the context (Kutnick and Baines 2001). As discussed in Chapter 2 the context dictates the impact of the mediating artefact. The disadvantage is the lack of control over other factors that influence the children's behaviour (MRC 2000). The imposition of an experimental approach also means the format is restricted throughout the study. No changes can be made to the software, or to the structure of the teaching. This contrasts with Action Research, which encourages variation as the focus is the outcome and not the individual factors that influence it (Cohen and Manion 1998, p.192).

The structure of the lessons imposed boundaries that could prevent the transfer of the awareness raised through using the DSGS v2 model (Crook 1999). These are listed below:

- The domain for raising awareness of group skills is restricted to ICT
- ICT lessons were held in the computer suite – "normal" lessons were held in the classroom
- The ICT lessons were unusual as they were given by the researcher not the class teacher

The study itself may not result in change due to an insufficient timescale. Adey and Shayer (1994) recommend two years as necessary to result in a change in cognitive and social understanding for older children if only used for one session a week. The changes identified through the quantitative and qualitative methods used, such as score results, questionnaire data, and discourse analysis, may not relate to the use of the DSGS v2 model. The initial puzzles could have influenced the children's sensitivity to working together and cloud the effect of the treatment (Cohen and Manion 1998, p.172). Furthermore, the information provided by the teachers is subjective and was not corroborated - other teachers may give different opinions. There are agreed criteria for believing children are working well together, e.g., body language, and amount of assistance required (see p.88), but this would be influenced by other knowledge, such as a history of that child and when they were observed.

Due to the limited number of participants caution is needed when drawing conclusions. It is a quasi-experimental setting – restricted to the schools volunteering and then by technical difficulties. A better design would have been to have groups where *all* or *none* of the children had used the DSGS v2 model. Another disadvantage is the children's age. Those permitted to

participate, Year 5, are also only just beginning to demonstrate the appropriate vocabulary for reflecting on the impact of their behaviour (Sharples 1999, Azmitia 1998).

The type of school also influenced the results. Daniels (2001b, p.152) argues that explicit teaching on how to collaborate through active interventions is insufficient unless it is embedded within the school's philosophy and practices. The teachers kept to their normal teaching methods as described in Study 1d throughout. They participated when requested, for example, assessing group skill usage, but the culture of the school was the transmission of curriculum subject content. This is demonstrated by their interest in the ICT quiz results and use of the course material in later years. This reinforced the boundaries between the training scheme and other subjects. A different outcome could occur if the teachers had shared the goal of the research, to improve group skill awareness, and been more active in its promotion (Galton and Williamson 1994, p.170).

8.2 Study 6a: ICT performance

On the basis of Study 2, the first hypothesis proposed that children who had worked in supported pairs would retain more of the information they had studied (hypothesis a, p.155). This would be supported by a better performance in the ICT quiz in Phase 3b.

8.2.1 Results

As there is no significant difference in their QCA scores a direct comparison can be made between Groups C and D (see Table 8.2). A comparison of means shown in Table 8.5 shows that compared to Group C the children in Group D: answered more questions, failed less questions, entered the correct answer with a smaller number of attempts, and solved the question quicker. Despite these trends, there is no significant difference using a t-test between any of the categories.

AVERAGE	GROUP C (n=25)	*Standard deviation*	GROUP D (n=26)	*Standard deviation*	*T-test*
QCA (out of 100)	45.46	*20.30*	48.71	*19.13*	*0.4964*
Number of questions answered	12.16	*3.33*	12.88	*3.54*	*0.4549*
Number of questions failed	0.84	*1.11*	0.69	*1.26*	*0.6585*
Number of attempts to solve question	3.28	*1.83*	3.00	*2.46*	*0.5349*
Time per question (seconds)	247	*77*	214	*82*	*0.1408*

Table 8.5: ICT quiz scores

8.2.2 Discussion

Although the results cannot be extrapolated given the small sample size, the trends could indicate that using the software does result in a marginal gain in the information retained, despite the fact that less time was spent covering the material to fulfil the GST requirements as described in Section 7.3. Alternatively, the trend may be due to the initial higher ability of children in Group D, or as a result of the individual's personality, experience and personal style of working. These have an equal or greater impact on task performance compared with the knowledge gained through working in a group (Mercer 1994, Crook 1998, Study 1d).

Assuming the DSGS v2 model influenced the information retained in Group D, several reasons could explain why it resulted in less task performance improvement than Study 2. The first is the attitude of the children. In Study 2 the task was seen as a puzzle, in Study 6a the focus is on the ICT domain knowledge. The presentation as a quiz may have influenced the attitude of the children towards it. Girls respond better to tasks that could be interpreted as tests (Light and Littleton 1999d). Another explanation may be the timescale. The children had to retain the information over a longer period and the delay could impact the post-test performance (Howe 1991). The TGT itself could influence outcome. The Escape from the Forbidden Forest task was a joint activity–joint outcome task that required participation from both children. The dual-key control constrained how the group interacted with the computer and ensured some communication and responsibility towards the completion of the task. In Aliens reflection, responsibility and communication are encouraged only within the GST. The design proposed that this would influence the use of these attributes in the TGT, as this is equivalent to encouraging reflection-in-action (Schön 1987). These prompts though, did not constrain how the group worked while doing the TGT. It would be possible for one child to dominate, or "freeload", throughout the task. The software encourages children to think about group skill usage but cannot force the children to negotiate roles, take turns etc. The differing levels of involvement could have influenced the amount of task information retained.

These results support Study 3a's finding that the compulsory self-assessment does not appear to be detrimental to the overall task performance.

8.3 Study 6b: Puzzles - task performance

From Study 2 it was hypothesised that those who had used the training scheme based on DSGS v2 model would complete a set of group puzzles more efficiently than the control groups (hypothesis b, p.155). That is, they would have solved the puzzle in fewer moves, with fewer errors, less assistance, and with a longer average time before making a move. These relate back to the skills the software was meant to develop. Longer pauses could mean more discussion, hence improved communication, or that reflection occurred.

The distribution of pauses could indicate the style of working. If the group agreed a plan and then executed it, there would be a long pause followed by short ones. Groups that plan before each move would have pauses of similar length (Issroff 1999). It is recognised the time recorded between moves does *not* necessarily represent time on the TGT. In some cases, pauses relate to arguments within the group, or being distracted by others. This cannot be illustrated through a numerical analysis. Attempting to complete the task without help suggests a degree of responsibility towards developing an individual and group understanding. The ability to work without referral to the teacher or waiting passively for the answer is a factor used to assess group skill usage (Study 2).

8.3.1 Results

In this section, the performance at each of the four puzzles in Phase 1 and 3a is given. The Restarts value in the tables in this chapter is the number of times the group began the task again. This could be due to errors or choice. The Assistance term at the bottom of each table represents the number of times the researcher helped the group. In the Return to Earth puzzle this includes the number of times the clues were viewed. In Group D all the children had used the training scheme based on the DSGS v2 model. In Group C all the children came from the control set. In Group M there was a mix of children, some had used the scheme and others had not.

Table 8.6 shows the Statue puzzle performance. The only significant difference between the task performance in Phase 1 and 3a is the decrease in the number of moves ($p>0.1$ using a t-test). There is a decrease, however, in the number of moves and time taken to make a move from Group C to Group M and then to Group D, with all the groups from Phase 3a taking less moves and assistance, and less or an equal time per move than in Phase 1. Group D took longer overall,

but unlike Group C, they had to restart. Group D also needed more help than the Group M, but less than Group C.

AVERAGE	PHASE 1		PHASE 3A					
	All (5 groups)	*Standard deviation*	All (5 groups)	*Standard deviation*	Group C (1 group)	Group M (3 groups)	*Standard deviation*	Group D (1 group)
QCA score (%)	49.78	*4.53*	45.77	*3.25*	48.57	46.59	*3.19*	41.00
Total moves	74.67	*2.12*	47.60	*10.97*	65.00	43.50	*9.19*	40.00
Time between move	0:00:09	*0:00:03*	0:00:12	*0:00:05*	0:00:09	0:00:10	*0:00:01*	0:00:20
Total time	0:10:55	*0:04:12*	0:09:24	*0:02:43*	0:09:16	0:06:50	*0:00:40*	0:13:00
Restarts	0.67	*0.00*	0.80	*0.45*	0.00	1.00	*0.00*	1.00
Assistance	14.67	*11.31*	6.20	*5.22*	13.00	4.00	*4.24*	9.00

Table 8.6: Statue puzzle performance

This contrasts with the trends in performance at the Frogs puzzle shown in Table 8.7. The only significant difference between Phase 1 and 3a is the increase in the number of moves ($p>0.1$ using a t-test). Group D took most moves, needed most help and had the shortest pause between moves. They had more restarts than Group C, who performed the task in fewest moves, and had the longest time between moves.

AVERAGE	PHASE 1		PHASE 3A					
	All (5 groups)	*Standard deviation*	All (5 groups)	*Standard deviation*	Group C (1 group)	Group M (3 groups)	*Standard deviation*	Group D (1 group)
QCA score (%)	47.40	*2.66*	47.57	*0.71*	48.13	47.78	*0.24*	46.40
Total moves	72.33	*9.81*	124.40	*57.42*	49.00	127.67	*39.80*	190.00
Time between moves	0:00:10	*0:00:07*	0:00:10	*0:00:06*	0:00:20	0:00:08	*0:00:03*	0:00:05
Total time	0:11:03	*0:06:16*	0:16:30	*0:04:52*	0:16:05	0:17:04	*0:06:47*	0:15:12
Restarts	9.67	*4.04*	9.00	*6.44*	1.00	12.00	*6.08*	8.00
Assistance	19.33	*5.77*	20.60	*11.76*	23.00	14.00	*7.55*	38.00

Table 8.7: Frog puzzle performance

The number of pairs who completed the Petshop puzzle rose from 13 out of 15 in Phase 1 to 14 out of 14 in Phase 3a. The results are shown in Table 8.8. There was no significant difference between the performance in Phase 1 and the combined performance of the three groups in Phase 3a. The only significant difference between Group C and Phase 1 was the decrease in the average time to make a move ($p>0.1$ using a t-test). On average Group D performed the task in the least

moves, had the longest average time between moves, took the longest to read the instructions, made the least errors and had most pauses greater than 11 seconds. Group C performed the puzzle in the most moves, took the shortest average time between moves, were the fastest to read the instructions, made most errors and had most assistance. There is no significant difference between these values in Group C and Group D using a t-test.

	PHASE 1		**PHASE 3A**					
AVERAGE	All who finished (13 of 15 pairs)	*Standard deviation*	Group C (5 pairs)	*Standard deviation*	Group M (4 pairs)	*Standard deviation*	Group D (5 pairs)	*Standard deviation*
QCA score (%)	48.04	*17.11*	48.07	*21.75*	54.27	*29.58*	46.48	*18.17*
No of moves	84.85	*53.97*	113.86	*93.96*	90.50	*111.02*	78.80	*58.47*
Time to read instructions	00:00:47	*00:00:26*	00:00:30	*00:00:14*	00:00:58	*00:00:35*	00:01:06	*00:00:52*
Time between moves	00:00:08	*00:00:03*	00:00:05	*00:00:01*	00:00:06	*00:00:01*	00:00:07	*00:00:03*
Total time	00:10:23	*00:04:56*	00:08:43	*00:04:52*	00:07:55	*00:09:21*	00:08:10	*00:05:08*
Number of pauses in five minutes								
5s - 10s	10.36	*2.90*	9.53	*2.84*	14.23	*1.63*	9.86	*5.08*
11s - 25s	6.26	*3.76*	4.69	*2.39*	3.64	*0.29*	5.04	*0.98*
> 25s	0.91	*0.79*	0.63	*0.62*	0.17	*0.24*	1.20	*1.46*
No of errors	21.15	*14.37*	24.14	*18.99*	17.00	*22.63*	15.60	*13.33*
Assistance	2.77	*2.05*	3.00	*2.45*	0.50	*0.71*	1.80	*1.30*

Table 8.8: Petshop puzzle performance with 4 tanks

Four out of 14 pairs completed the Return to Earth puzzle in Phase 1. In Phase 3a, one pair in Group D did not complete it, and another pair from Group D was not recorded due to technical difficulties. The results are shown in Table 8.9. There was no significant difference between the performance in Phase 1 and Group C, or Phase 1 and the combined groups in Phase 3a. On average Group C performed the task in least moves, took longest to read the instructions, had the longest average time between moves and most pauses greater than 11 seconds. Group M performed the puzzle in the most moves, were the fastest at reading the instructions, took the shortest average time between moves, and made most errors. The only significant difference between Group C and D is that of overall time. Group D are slightly faster ($p>0.1$ using a t-test).

AVERAGE	PHASE 1		PHASE 3A					
	All who finished (4 of 14 pairs)	*Standard deviation*	Group C (5 pairs)	*Standard deviation*	Group M (5 pairs)	*Standard deviation*	Group D (3 pairs)	*Standard deviation*
QCA score (%)	49.72	*25.96*	39.71	*15.88*	48.35	*17.46*	50.48	*21.57*
No of moves	61.75	*37.85*	60.50	*10.61*	71.89	*29.47*	62.00	*29.70*
Time to read instructions	0:01:10	*0:00:30*	00:01:26	*00:00:50*	0:01:03	*0:00:27*	00:01:17	*00:00:24*
Time between moves	00:00:22	*00:00:06*	0:00:22	*0:00:03*	00:00:15	*00:00:02*	00:00:20	*00:00:11*
Total time	0:19:57	*0:06:09*	0:22:25	*0:01:07*	00:17:03	*00:05:54*	0:17:59	*0:01:40*
Number of pauses in five minutes								
5s - 10s	5.34	*3.57*	3.76	*1.39*	10.31	*1.88*	6.85	*3.78*
11s - 25s	8.71	*2.51*	7.91	*0.39*	6.66	*1.87*	5.25	*2.26*
> 25s	2.25	*1.32*	3.26	*0.95*	2.20	*0.85*	2.81	*2.29*
No of errors	15.75	*10.75*	14.00	*1.41*	16.67	*6.10*	13.00	*9.90*
Assistance	5.5	*3.77*	6.00	*0.00*	5.89	*2.62*	8.00	*1.41*

Table 8.9: Return to Earth puzzle performance

Performing the Statue and Petshop puzzle in the same session does not appear to affect the task performance. After completing the Statue, one pair went on to do the Petshop puzzle in the least number of moves, whilst another pair from the same group took the second greatest number of moves. Similarly changing the order of the puzzles had no impact. A group of six actually took more moves than the average to complete the Statue, despite all three pairs successfully completing the Petshop puzzle immediately before.

8.3.2 Discussion

The improvement in performance in the Statue puzzle, and the worsening in the Frogs puzzle may be related to the groups rather than the puzzles. Although there is no significant difference in ability the half who did the Statue puzzle first may be more able at puzzles, and so finished the tasks in fewer moves in both cases. This does not explain the performance in the puzzles for pairs where there was no significant improvement. This suggests it could relate to larger group sizes and the friendship groupings and competition within them (Pauli and Reusser 1997, Linehan and McCarthy 2001, Study 1d).

These findings suggest that awareness raised of group skills using the DSGS v2 model, in the form of the Aliens software, does *not* transfer to other group activities. This includes near transfer tasks, i.e., the Petshop Puzzle and Return to Earth, which are both computer-based

activities for pairs. For the Frog and the Return to Earth puzzles the hypothesis that using the DSGS v2 model results in a more efficient outcome is contradicted. In these puzzles on average Group D took most moves and had the shortest pauses between moves. In no puzzle was the hypothesis that children in Group D would request the minimum assistance fulfilled. This implies the children were not taking responsibility for the outcome. Neither was reflection, for example, planning moves ahead, observed. In groups of six, children in Group D never took least restarts. Finally, communication seemed erratic. Group D took the longest time between moves doing the Statue and Petshop puzzles, and shortest time for the Frog and Return to Earth puzzles. This again suggests the context has most impact on the outcome.

Although using the model does not result in increased time between moves, there is a relation between this and the number of moves taken to solve the puzzle. In three of the four puzzles, those with the longest average time between moves took least moves. While in all cases those with the shortest average time between moves took most moves to complete the puzzle. This supports the suggestion that the length of pauses reflects task performance (Wegerif 1996b). In the puzzles for pairs, the group with the least moves had most pauses greater than 11 seconds, although poor performances can also have long pauses. In the Petshop puzzle, for example, those in Group C had more long pauses than Group D, but had more errors, required more assistance, and took more moves. Longer pauses may be due to off task activities, such as arguing. They may, however, indicate the children are planning or checking for agreement. This implies the claim that reflection and communication are necessary attributes to work well as a group is correct. The study does not validate the third attribute of responsibility, as improved performance does not correspond with fewer requests for help.

The lack of improvement when performing versions of the Tower of Hanoi puzzle consecutively supports the findings from the design stage and of Pierce and Gholson (1994). They suggest surface similarities impact puzzle recognition and hence task performance.

8.4 Study 6c: Puzzles - discourse analysis

The previous section showed using the DSGS v2 model had little impact on task performance. In this section, the impact on language is investigated (hypothesis c, p.156). This section is divided into two parts. The first part is a quantitative discourse analysis. Coding words or speech acts can be ambiguous; the function of words is dependent upon the context (Draper and Anderson 1991). To overcome this an analysis was made of keyword usage rather than the occurrences of words in a transcript (Mercer and Wegerif 1999). The second part is a combination of qualitative and quantitative discourse analysis. Case studies are presented on two girls: one who improved in group skill usage according to her teacher, and another who stayed the same. The talk is categorised as being exploratory, cumulative or disputational (Mercer 1994). These were chosen as the training scheme promoted exploratory talk through encouraging discussions containing planning, justifications, turn taking, and role taking. This analysis would corroborate the quantitative analysis, by showing whether the principles within the DSGS v2 model have transferred to other activities. It also investigates the theory that the type of talk is related to task performance (Wegerif 1996b, 2001).

8.4.1 Quantitative analysis

It was hypothesised that children in Group C would use similar language in Phases 1 and 3a, and children in Group D would use more exploratory talk than Group C in Phase 3a. This would be because such behaviour had been modelled by the Aliens software and appropriated. Exploratory talk includes asking task-focused questions, justifications, challenging, considering multiple positions, drawing opinions from all and reaching an agreement. This analysis focuses on participation, statement length and justifications. A larger improvement was expected in the near transfer task - the puzzles for pairs, compared to the far transfer task – the puzzles for six. To test if the model had been internalised, mixed groups of six were compared. It was predicted that children from Group D would use more exploratory talk features, and this would lead to improved talk by those from Group C.

8.4.2 Method

An analysis was made of the complete transcripts for seven pairs and four groups of six who had worked together to complete the puzzles in Phases 1 and 3a. It used the key linguistic feature of justifications suggested by Wegerif and Mercer (Wegerif 1996b, Mercer and Wegerif 1999). The number of occurrences of "because", "cos", "if" or "so" used in an exploratory sense were

recorded. This meant: "No you can't because 4's heavier than 3 and 5's heavier then urm 4…", was included, but: "if we ever complete this I'll be like bedazzled", was not. The DSGS v2 model was designed to encourage turn taking and listening to your partner. To confirm this the number of interruptions and occurrences of simultaneous speech were also recorded. Finally, as it has been suggested longer statements reflect exploratory talk this has also been calculated (Mercer *et al.* 1999).

To allow for comparison the proportion of speech acts are presented over a five-minute interval.

Three pairs from Group C were compared to four pairs from Group D performing the puzzle for pairs. Four groups of six were analysed, they comprised of:

GROUP	CONDITION
1	All from Group D (had used the training scheme based on the DSGS v2 model)
2	All from Group C (the control group)
3	All from Group D but coming from two groups consisting of three children from Group C and three from Group D
4	All from Group C but coming from two groups consisting of three children from Group C and three from Group D

8.4.3 Results

As shown in Table 8.10 in all pairs there was in increase in justifications for the puzzle moves, and a reduction in simultaneous speaking between Phases 1 and 3a. The trend is for children in Group C and D to display less simultaneous speaking and give more justifications in Phase 3a compared to Phase 1. Group C children interrupt more but have a longer average statement length and an increase in the number of words spoken. In contrast, children in Group D tend to interrupt less, but use fewer words and shorter statements. There was no significant change in performance between the discourse *within* groups during Phases 1 and 3a. An analysis between groups in Phase 1 shows no significant difference between Groups C and D. In Phase 3a, however, children in Group C significantly interrupt more; speak more; and use longer statements than those in Group D.

IN FIVE MINUTES THE AVERAGE NUMBER OF:	3 PAIRS FROM GROUP C			4 PAIRS FROM GROUP D			GROUP C AND D	
	Phase 1	Phase 3a	*T-test*	Phase 1	Phase 3a	*T-test*	*Phase 1 (T-test)*	*Phase 3a (T-test)*
justifications	3.45	5.22	*0.513*	4.08	4.13	*0.922*	*0.801*	*0.728*
interruptions	3.97	5.44	*0.326*	2.12	1.26	*0.265*	*0.326*	*0.011*
simultaneous speaking	0.90	0.44	*0.671*	2.72	2.03	*0.528*	*0.109*	*0.417*
words	397	464	*0.235*	437	304	*0.328*	*0.754*	*0.064*
average statement length	8.2	10.4	*0.215*	17.6	5.8	*0.273*	*0.325*	*0.006*

Table 8.10: Comparison of exchanges in dyads

To determine ability at puzzles the performance of the pairs within the groups were examined. Two of the three pairs in Group C were better than average in terms of moves taken to solve the puzzles in both Phase 1 *and* 3a. While the children in Group D appeared more erratic; two pairs did better than the average in Phase 1, but worse in Phase 3a, and the remainder did worse in Phase 1, but were better than average in Phase 3a.

Table 8.11 shows the results of the discourse analysis for groups of six.

IN FIVE MINUTES THE AVERAGE NUMBER OF:	GROUP											
	1 (GROUP D)			2 (GROUP C)			3 (GROUP D)			4 (GROUP C)		
	Phase 1	Phase 3a	*T-test*	Phase 1	Phase 3a	*T-test*	Phase 1	Phase 3a	*T-test*	Phase 1	Phase 3a	*T-test*
justifications	3.78	7.09	*0.368*	6.74	5.50	*0.614*	4.15	10.82	*0.366*	3.71	5.32	*0.672*
interruptions	8.54	15.75	*0.188*	17.89	7.25	*0.150*	7.51	2.75	*0.069*	13.85	10.71	*0.723*
simultaneous speaking	11.33	5.77	*0.057*	12.19	6.75	*0.176*	32.78	14.61	*0.172*	25.22	20.49	*0.674*
Words spoken by children	460	526		665	535		568	529		536	495	
AVERAGE STATEMENT LENGTH												
Overall	5.82	4.35	*0.018*	5.46	7.19	*0.118*	5.59	5.24	*0.790*	4.67	4.70	*0.979*
NUMBER OF STATEMENTS PER CHILD:												
Child 1	23.05	27.82		2.59	5.25		5.46	2.08		16.78	17.71	
Child 2	15.50	37.27		24.37	9.25		8.58	6.25		5.46	1.04	
Child 3	6.95	9.19		23.85	14.75		23.41	22.92		23.02	21.88	
Child 4	9.74	2.62		14.26	9.75		21.59	13.68		4.39	5.13	
Child 5	0.99	15.49		8.56	10.00		33.66	27.35		43.17	18.80	
Child 6	17.28	28.08		30.60	21.25		8.78	3.42		9.51	23.08	

Table 8.11: Comparison of interactions within groups

From Table 8.11 an analysis *within* a group's performance from Phase 1 to Phase 3a shows that all groups except Group 2 increase the number of justifications. Over this period Group 1 was the only one with an increase in interruptions and words spoken, the former almost doubled. All the groups displayed less simultaneous speaking in Phase 3a than Phase 1. After using the model, the average statement length decreased for Groups 1 and 3. This was a significant drop for Group 1, from 5.82 to 4.35 words ($p>0.02$ using a t-test). Without exposure to the DSGS v2 model the overall average statement length of Groups 2 and 4 increased. A comparison of the number of statements indicates that although speaking in shorter statements all but one of the children in Group 1 spoke more in Phase 3a than Phase 1. Although also exposed to the model all the children in Group 3 spoke less frequently after its use. While for children not exposed to the model in Group 2 only two children spoke more but the average statement length increased, and in Group 4 the statement length remained similar with half the children displaying an increase in the number of statements and the others a decrease.

When comparisons were made *between* Groups 1 and 2 in Phases 1 and 3a no significant differences were found in the number of justifications, interruptions, simultaneous speaking and average statement length. Nor were there any significant differences in the occurrences of these linguistic acts between those in Groups 3 and 4 when compared at each phase. This lack of significant difference in these categories was also true when comparing a combination of all the children who had used the DSGS v2 model (Groups 1 and 3) and all the children from the control group (Groups 2 and 4).

Children in Group 1 were asked if they could remember the rules for working together after the activity. They looked surprised and then upset as they admitted they had forgotten them when performing the puzzle. All the groups wished to know whether they had performed the puzzle better than the other groups. Regardless of group the quieter children explained their silence by being more interested in having the puzzle performed well, consequently they were willing to let those they believed were the more able take charge.

In terms of puzzle ability Group 1 were more able than Group 2. They took fewer than the average number of moves to solve the puzzles in Phases 1 and 3a while Group 2 took more moves than the average in both cases.

8.4.4 Discussion

The results do not support the hypothesis in this situation. The predicted improvement in language use when comparing children working as pairs and in groups of six between Group D and Group C did not occur between Phases 1 and 3a. A comparison within Group D showed the only significant difference was a decrease in the length of statements when working in a group of six. It was hypothesised statement length would increase. A comparison within Group C showed no significant change in discourse style between Phases 1 and 3a.

The increase in justifications is not comparable to that observed when talk lessons were integrated into the classroom (Mercer and Wegerif 1999). In all cases, the number of interruptions increases if the number of words spoken increases overall, and reduces if there is a reduction in the total number of words. Comparing the average statement length it was found to be shorter in Group D in Phase 3a than Phase 1, and as stated significantly so when working in groups of six. In contrast, the statement length increased in Group C, and a comparison in Phase 3a showed it was significantly longer than Group D when working in pairs. It is possible this is related to the ability at puzzles. More children in Group C were consistently better at puzzles than pairs in Group D, and those in Group 1 who all came from Group D were better at puzzles than those in Group 2 who all came from Group C. If it is assumed, however, statement length reflects exploratory talk (Mercer *et al.* 1999) then the model is detrimental. This implies that requiring self-assessment of group skill usage and modelling exploratory talk through written prompts when performing a group task is *not* a suitable method for raising awareness of group skills in this age. Sessions once a week do not result in the children appearing to have internalised the group skills discussed in the Aliens sessions and integrated them into all group activities.

Study 3b suggested that children were capable of providing justifications and reaching a consensus. This was in the presence of the researcher. Without reminders, the children appear to focus on the puzzle, which reflects the focus on the TGT in Studies 1, 2, 3a and 5. This mirrors the object of the classroom environment (Kuutti 1996), in this case the transmission of domain knowledge. The children acknowledging they forgot the rules for effective group work when given a task in another environment reinforces this. This finding re-emphasises the need for reminders to be given at the start of each session (Lin 2001, Kramarski and Mevarech 1997, Salomon 1984).

It should also be remembered that the exposure to the mediating artefact, in this case the Aliens software, is not the only factor affecting discourse. The desired outcome could have been perceived as puzzle completion and *not* having individual task understanding, i.e., they could have been treated as co-operative (individual activities-joint outcome) tasks. This contrasts with The Escape from the Forbidden Forest task in Study 2 where there was a responsibility towards ensuring all understood. This may explain why the quieter and less able allowed the others to dictate moves. In addition, the concept of activity settings shows the age, friendship groupings, type of task etc., are influential (Pauli and Reusser 1997, Linehan and McCarthy 2001, Mercer 1994).

8.4.5 Qualitative analysis

The previous section suggests that the training scheme based on the DSGS v2 model has little impact on the type of exchanges. Given the opinions of the teacher from the previous chapter on feedback this result is unsurprising. Only some children who used the DSGS v2 model improved in terms of group skill usage. This could explain the overall lack of difference in the number of justifications, interruptions etc. To corroborate the teacher's opinions about improvement in group skill usage, and confirm whether exploratory talk was displayed by those that had improved, the discourse of two girls was analysed. According to their teachers Becky's use of group skills had improved a lot, while Daisy's had remained the same. Transcript extracts as well as quantifying exchanges are used to minimise the loss of information (Edwards and Westgate 1994a).

8.4.6 Method

The two girls were selected as representing differing outcomes from using the DSGS v2 model. Both had used the Aliens software for 18 weeks. They worked with the same partner in Phases 1 and 3a. Their teachers considered their partners as being good at using group skills. The girls had worked with other children during the exposure to the model in the ICT lessons. Becky's usage of group skills was initially poor but had improved a lot according to her teacher. Daisy's usage of group skills remained good throughout. (There are no cases in which using the DSGS v2 model led to a decrease in using group skills according to the teachers.) A qualitative analysis was made of the interactions between the children in Phases 1 and 3a, and a quantitative analysis was made similar to that in Section 8.4.3.

The categories used were cumulative and exploratory (Mercer 1994). Cumulative talk is characterised by repetitions, confirmations and elaborations. The speakers build uncritically on the comments of the other. Exploratory talk has the reasoning visibly displayed, challenges to statements are justified and alternative hypothesis offered. The partners offered constructive criticism.

8.4.7 Results

The two case studies are described in turn. Discrepancies exist between the number of statements between the pair, as remarks addressed to other children or the researcher are ignored.

CASE STUDY 1: BECKY

Becky was less academically able than her partner James. She achieved a mean of 32.33% in her QCA examinations while he achieved 81.42%. In Phase 1 they completed the Petshop puzzle in 68 moves over 4:26 minutes - the average was 77 in 10:23 minutes. James took charge throughout the activity. In the 21 minutes they worked together he had control of the mouse for the majority of the time and said 572 words in 8 monologues. This compares to Becky's 35 words in 5 statements; she asked two questions, made two suggestions that were ignored or interrupted, and announced the task was finished. James ability and control over the task may explain the above average performance and Becky's lower ability could explain her diffidence and allowing her partner to take control. Another explanation is the relationship between the children; friends, for example, are more likely to talk (Puntambekar 1995). An example of their exchange is given in Table 8.12.

	NAME	DIALOGUE	ACTIONS
1	James	This time you try to do it again with 4 tanks	Gives Becky the mouse
2		No that one	James takes the mouse back
3			Becky takes mouse back as James lets go to turn round
4		No up there	James points to screen - hand hovers by mouse Becky relinquishes control
5		It's not the smallest	James moves fish
6			James lets go of mouse and Becky takes again
7		No	Becky lets go of mouse
8		Watch this	James takes
9		I'll just show you again	Becky put hands in lap when not with mouse
10		You do the rest	Becky takes mouse
11		Move that one there	James points to screen

Table 8.12: Extract from transcript in Phase 1

Given the disproportionate amount of time spent talking by James classifying the conversation as cumulative, disputational or exploratory is problematic. James takes a highly didactic approach, possibly copying the approach taken by his teacher (Study 1d), and Becky acts passively. When Becky had the mouse James instructed her, as in lines 1, 4 and 11. The exchange is almost peer tutoring. James appears to have taken responsibility for teaching; offering explanations, but not checking Becky's understanding (line 5), and volunteering to demonstrate (line 9). The extract supports the teacher's claim that Becky was "very bad" at working in groups.

In Phase 3a Becky and James worked together for 25 minutes. Unlike Phase 1 their task performance was below average, they took 83 moves in 15:48 minutes - the average was 69 in 18:01 minutes, to solve the Return to Earth puzzle. As before James spoke most, uttering 733 words in 120 statements averaging 6.66 words. However, Becky's contribution increased: she said 387 words in 85 statements with an average length of 4.55 words. An extract from the transcript is given in Table 8.13.

	NAME	DIALOGUE	ACTIONS
1	Becky	No 'cos then he'll get eaten, the policeman will be	
2	James	Oh yeah	
3		2 aliens go up	Sending the policemen up as speaks
4		You've got to send 2 aliens then 1 back…	Becky is watching
5		2 aliens there and 1 back?	Becky stops twiddling with traffic lights
6	Becky	Yes, there's 1 up there so send 2 aliens.	James moves pieces from shuttle to Zog, then an alien back to Zog
7		Now send 2 policemen	
8	James	No, 'cos that's where we got stuck	
9		Urm - we could try it …and then send…	Fades out
10	Becky	An alien back and a policeman	
11	James	'Cos if you send those 2 back you've got to send a policeman with it	Uses mouse as a pointer
12	Becky	Urrr… an alien and a policeman	James sends them to the shuttle

Table 8.13: Extract from transcript in Phase 3a

Categorising the talk as cumulative or exploratory talk is problematic as both are demonstrated. In this extract, the children do plan moves and partially justify their statements: James justifies why he is rejecting Becky's plan of sending 2 policemen in line 8, Becky rejects a move in line 1, as the policemen will be eaten. In line 12 however, there is no justification for her suggestion.

The fact that James performs the move suggested can be seen as cumulative. Line 5 is ambiguous. It may illustrate cumulative talk and be intended as a confirmation, or exploratory acting as a check for agreement.

Table 8.14 is a qualitative analysis showing an increase in the number of justifications made by both children. Becky's participation increases almost 17 fold. Her average statement length decreases and James interrupts her more. The children never spoke simultaneously.

IN A FIVE MINUTE PERIOD:	BECKY		JAMES	
	PHASE 1	PHASE 3A	PHASE 1	PHASE 3A
Average statement length	7.00	4.55	71.50	6.66
Number of statements	1.00	16.89	2.00	22.25
Number of justifications	0.00	1.79	1.00	2.78
Number of interruptions	0.25	0.20	0.25	1.39
Number of times speaking simultaneously	0.00	0.00	0.00	0.00

Table 8.14: Quantitative analysis of interactions

The improvement in Becky's use of group skills noticed by the teacher and from her discourse may be attributable in part to the training scheme. This is in conjunction with variables outside of the study's control, such as maturation or other lessons (Cohen and Manion 1998, pp.170-171). It is also possible that Becky developed an enthusiasm for puzzle solving, she found the Escape To Earth puzzle more stimulating, or James had more influence on her behaviour (as demonstrated in line 5 of Table 8.13 he now asks questions instead of giving instructions). These explanations, however, do not explain her improved group skill usage noted outside the study environment.

CASE STUDY 2: DAISY

Section 8.4.1 shows some children who used the DSGS v2 model regressed in terms of conversation displayed. An example is Daisy. Daisy averaged 45.42% in the QCA exams, and worked with Holly, who scored 50.83%. As shown in Table 8.15 the talk during the Return to Earth puzzle in Phase 1 was a combination of cumulative and exploratory talk. Lines 7 to 10 and 14 to 17 are social exchanges, some planning occurs in lines 11 and 12, and explanations are given to justify opinions in lines 1 and 13. The girls also share using the mouse and appear interested in the moves made by their partner. Despite remaining on task and working together for almost 17 minutes, they failed to complete the puzzle.

	NAME	DIALOGUE	ACTIONS
1	Daisy	YES! Now we have to take one alien back down - no you can't do that as it can't fly by itself	Hollie tried to click on blast off with empty shuttle
2	Hollie	Ohhhhh	
3	Daisy	What? What did you put? What did you do?	
4	Hollie	I just took 1 alien back down with me	
5	Daisy	Hold on - you - you need to put |	
6			Hollie briefly asks other group where they are - Daisy continues
7	Daisy	… go on just do it - ohhh	At error puts head in hands and sprawls over table
8	Hollie	I'm taking him back down	Hollie takes mouse
9	Daisy	this had better work - blast off	
10	Both	Yes!	
11	Daisy	Which one shall we take?	Daisy picks up mouse, Hollie let go to raise hands
12	Hollie	Take take 1 policeman and 1 alien up - yes this is going to work now |	
13	Daisy	No because they'll be they'll be more aliens - you need to take 2 policemen - cos then it's equal - you can only fit 2 people in… … yeah we know that - where's that other policeman? What?	last bit reading from screen, assume trying to get 3 in shuttle
14	Hollie	<laughs>	
15	Daisy	Where's the other policeman gone?	
16	Hollie	He must must have ate him	Hollie points to screen
17	Daisy	Oh we've lost it - we've lost it!	

Table 8.15: Extract from transcript in Phase 1

This variation of exchanges contrasts to the dialogue from Phase 3a illustrated in Table 8.16. Daisy and Hollie took 63 moves in 7:30 minutes to complete the Petshop puzzle, the average was 98 in 8:24 minutes. They worked together for just over 16 minutes with Daisy having the mouse throughout. Hollie has little interest in the task and is talking about earrings until line 12. Although the discourse is cumulative in nature, both girls rarely simultaneously focus on the task, nor do they ask questions, offer full explanations, or take responsibility for their partner's understanding. As in the case of Becky this may be an influence of age, Hollie may have become more interested in appearance than puzzle solving and is distracting Daisy, although her teacher noted no change in her use of group skills.

	NAME	DIALOGUE	ACTIONS
1	Hollie	What sort of earring clips?	She is fiddling with own earrings, then feels the back of Daisy's earrings
2	Daisy	And that into that bowl	Talking to screen ignoring Hollie
3	Hollie	What sort?	She leans round so she can look at the back of Daisy's ears - Daisy still working on puzzle
4	Daisy	We've got to get that - oh god	Twiddles with lip while Hollie still points at her ears
5	Hollie	{Can I see?}	Still referring to earrings
6	Daisy	{So we've got to put one into there}	Daisy stops to twiddle with earrings
7	Daisy	Yeah	Checks both ears
8	Hollie	I'm going to get one put up there	Points to top of ear
9	Daisy	I did have one stuck on there but it came off	Points to top of ear
10	Hollie	Everyone knows about you and …	Still twiddling with her ear, Daisy is looking round room
11	Daisy	What happens if we… oh dammit I forgot if we… ohhhh	Trails off and starts working on problem again, Hollie also sits forward
12	Hollie	She said we could move that one	No gestures just looking
13	Daisy	And if we move that one…	
14	Hollie	Yeah	
15	Daisy	I've got it I've got it … oh no I can't do that	
16	Hollie	Move that	Points to screen
17	Daisy	Oh no, I'm just going to have to do that	Continues without waiting for a response

Table 8.16: Extract from transcript in Phase 3a

This is reflected in a qualitative analysis of their scores in Table 8.17. The number of statements, justifications, and statement length for both girls decreases.

IN A FIVE MINUTE PERIOD:	DAISY		HOLLIE	
	PHASE 1	PHASE 3A	PHASE 1	PHASE 3A
Average statement length	14.20	7.96	8.24	4.39
Number of statements	24.75	20.57	17.09	3.58
Number of justifications	5.31	1.71	0.88	0.00
Number of interruptions	2.95	0.00	2.06	1.71
Number of times speaking simultaneously	3.24	2.57	3.24	2.57

Table 8.17: Quantitative analysis of interactions

8.4.8 Discussion

Using discourse analysis is problematic. It does not take into account the histories of the participants, friendship groupings, attitude towards task, etc. (Littleton 1999, Pauli and Reusser 1997). The case studies do corroborate the opinion of the teacher. Children who improve in group skill usage employ more justifications and participate more in group activities: those judged the same have a similar or worse interaction style.

In both cases, the better task performance involves less exploratory talk. Table 8.14 and Table 8.17 show a decrease in the number of justifications, and, apart from James who dominated the discussion in the winter, a decrease in average statement length with better task performance. This contradicts existing research which shows that such talk correlates to improved task performance (Mercer and Wegerif 1999). It does, however, support the finding that discussion does not represent individual understanding (Fisher 1993).

As Becky is considered to have improved in group skill usage outside the study it can be assumed that her improvement in the activities is not totally due to the Return to Earth puzzle. This, for both girls, generated more justifications and longer statements. It suggests that the DSGS v2 model could have had some influence. It may not be the only factor. As discussed factors such as friendship pairings, partner's behaviour, even an increase in age could influence performance and the type of interaction. Salomon (1995) states software can afford new opportunities and support for collaborative learning, but whether the performance improves outside is dependent far more on other factors than technology. This is demonstrated by Daisy not being influenced by Aliens in the same way. It also suggests some children may be able to internalise appropriate behaviour through modelling of the type of exchanges necessary for effective group work. This could relate to metacognitive ability and focusing on two different tasks simultaneously.

8.5 Study 6d: The teachers' questionnaire

It was hypothesised the teachers would observe an improvement in the group skill usage within those children who had used the DSGS v2 model compared to the control group (hypothesis d, p.156).

8.5.1 Results

In the opinion of the teachers, there was a significant improvement in the ability to work in a group over the period of the study ($p>0.001$ using chi square). This was true whether or not the child had used the DSGS v2 model. At the end of the study there was no significant difference in group skill usage between the control group, Group C, and those who had used the DSGS v2 model, Group D. This is shown in Table 8.18.

HOW MUCH HAVE THEY IMPROVED SINCE JANUARY?										
	A lot worse	A lot worse (%)	Worse	Worse (%)	Stayed same	Stayed same (%)	Better	Better (%)	A lot better	A lot better (%)
Group C	0	0.00	0	0.00	12	40.00	14	46.67	4	13.33
Group D	0	0.00	0	0.00	11	39.29	15	53.57	2	7.14
HOW GOOD ARE THEY AT WORKING TOGETHER?										
	Very bad	Very bad (%)	Bad	Bad (%)	OK	OK (%)	Good	Good (%)	Very good	Very good (%)
Group C	0	0.00	1	3.33	16	53.33	12	40.00	1	3.33
Group D	0	0.00	3	10.71	12	42.86	13	46.43	0	0.00

Table 8.18: Opinions of teachers on group skill usage

8.5.2 Discussion

The improvement in using group skills regardless of exposure to the DSGS v2 model suggests such skills are developed outside the computer suite. The use of group skills did not transfer from the ICT lessons into everyday activities. This could be due to the medium used to implement the model. As stated in Section 8.1.4 a computer-based activity is separate to those within the classroom. The expectation of using a computer is the gain of the skills specified by the National Curriculum, with any feedback relating to quantifiable skills, such as an understanding of mathematical principles. These results support Crook (1991); working at a computer can be seen as separate to more traditional instructional lessons. The skills required are assumed relevant only to that environment (Salomon 1995).

The improvement may also be age related. In the six months the child's linguistic ability may have increased sufficiently for their metacognitive awareness to improve. This would allow the children to work more effectively in groups as they can reflect on the tacit script given for working together and amend it.

One limitation discussed in Section 8.1.4 was reliability. It was suggested that despite similarity in opinions and criteria used when reviewing videos of children working together (see Section 4.2.7, p.88), other knowledge would influence the teacher's opinions. This includes the history of the child, the recent behaviour, and recalling the behaviour of the child at the start of the study. The process of observation itself may impact the behaviour displayed (Biott and Easen 1994b). These opinions, however, are corroborated through the case studies in Section 8.4. Improvement in group skill usage corresponds to more turn taking, justifications, etc., when performing a group task. Despite giving little credence to the opinion of children (Study 1d) the teachers opinions are further supported by the correlation between their views, and that of the children assessing group skill usage in Section 7.5. This suggests the questionnaire is reliable and that the Aliens software has not affected the attributes of reflection, responsibility, and communication. The boundaries of domain, class location, and teaching were not overcome.

8.6 Summary

This chapter describes the intervention study. Such an exercise was required as knowledge is constructed over time in a shared context. Two classes participated in the study; half used an implementation of the DSGS v2 model and half acted as a control group. The study took the form of pre-test, system use, post-test. The task performance was supplemented by the individual's performance on a quiz covering materials taught while using the system, and questionnaires on group skill usage completed by the teachers.

The quiz to test the information retained over the course of the 18 weeks showed a trend in which those who had used the model correctly recalled more of the information studied in less time. The differences were not as significant in Study 2. It was hypothesised this was due to the time the information had to be retained over, and the limited constraints on interactions within the task. The model prompts pupils to reflect on their interactions, which through appropriation impacts behaviour, rather than constraining the TGT behaviour itself.

The issue of transfer was not answered. It was hypothesised that children who used the model would be more aware of group skills and complete group tasks in fewer moves, with less help, and fewer errors. The improvement would be greater in tasks for pairs working at a computer. A relationship was found between time taken to make a move, and task performance in terms of move, however, it did not relate to exposure to the model. Using the model did not result in more longer pauses, indicating more reflection and communication, or fewer requests for help,

which would suggest taking responsibility for the tasks outcome. This could be to do with the groups, which were allocated not chosen. It could also be influenced by the context. The school focused on domain specific curriculum requirements. Although the method employed was related to Action Research, the teachers were not active members of the project. Their style of teaching was didactic with little reinforcement of group skills outside of the ICT lessons. The amount of group work done within the classroom environment did not increase during the system use phase, so there was little opportunity for the children to practise group skills in other activities.

A similar finding was made when examining the discourse. Children who used the Aliens software did not display more justifications, longer statements, less interruptions and simultaneous speaking because of having this type of discussion prompted during the ICT lessons. This implies that the model does not result in the appropriation of an awareness of the benefits of group skill usage. The focus is the TGT and not the GST. This was corroborated by the observations of the more silent children, who believed the task would be solved more efficiently if organised by the more able – regardless of their own understanding. A comparison of individuals supported the idea that it was context, rather than exposure to the model, that was influential. One girl used more exploratory talk structures, while another with equal exposure used less in the final phase. However, in these cases fewer moves were needed in puzzles where this style of exchange was not used.

The overall improvement in group skill usage recorded by teachers is corroborated by the case studies, and children's opinions. It shows that improvement is not related to using the software. Instead, as children approach the end of Year 5 for most their ability to work in groups improves.

These findings cannot be extrapolated given the type of school, small numbers and environmental limitations. These include the increased maturity of the children, the software being used outside of the classroom, the lessons run by the researcher, the model associated only with ICT, and the majority of time spent focused on the acquisition of curriculum specified domain knowledge. They suggest, however, that the behaviour of the majority of children aged 9 and 10 is *not* influenced by the model. The appropriation of feedback and prompting for exploratory discussion does not impact on future behaviour as the focus remains on the TGT rather than the GST. The teachers request for the lesson material rather than the DSGS v2 implementation reiterate this emphasis.

Chapter 9

REVIEW

This research investigated whether a training scheme to raise awareness of group skills when interacting face-to-face led to a change in the behaviour of children aged 9 and 10. The scheme was based on a model in which a classroom group activity consists of two tasks. The first is the teacher-given task (TGT). This relates to a National Curriculum domain that can be directly assessed, for example, creating a class book on animals around the world. The second focuses on general learning skills, including group skills such as listening, turn taking and ensuring understanding. This is the group skills task (GST). Group skill usage is harder to assess as no agreed criteria exist. In the training scheme the children are given a group activity in which the GST is supported by the Developing and Supporting Group Skills (DSGS) model. The approach is unique as:

- Raising awareness of the GST was designed to influence, *and* be influenced by, the TGT
- The feedback on group skill usage is generated from individual self-assessments made during and after the activity

This training scheme was developed incrementally. It incorporated research shown to be effective in encouraging group skills. Later studies corroborated and built upon earlier ones. With the evaluation methods chosen, children exposed to the final training scheme demonstrated: i) a higher individual cognitive ability compared to an equivalent control group – demonstrated by TGT performance and recall, and ii) the same group skill usage and task performance as the control group in later group activities. However, these results cannot be attributed solely to the mediating artefact, in this case the training scheme. The sociocultural and Activity Theory perspective taken in the design and analysis of the research placed equal emphasis on the context. In the initial studies this was demonstrated as differences in self-assessment of group skill usage when the medium (paper or software) and the environment (a classroom or laboratory-style setting, i.e., working with the researcher outside a standard lesson) were changed.

In addition to the direct impact of the training scheme on cognitive and social skills, this research has other implications. In the educational field the scheme has the potential to be used as a diagnostic tool: as a relationship was found between the self-assessments recorded and the child's observed ability at group work. From a technical stance, the research suggests these self-assessments and the time the children spent viewing the GST screens are sufficient to generate feedback about group skill usage that correlates with the opinion of a teacher. Finally, it supports the claim that the context and evaluation methods used influence the measured results of the training scheme.

This chapter re-examines the theoretical basis of the training scheme and discusses the implications of the research project. It is divided into four sections:

- The validity of the theoretical basis: a retrospective analysis of the four findings from the literature review that formed the basis for the research project
- The research project contributions: a discussion of the major contributions of this research to the current understanding of group work and computer supported collaborative learning (CSCL)
- Future work: this is divided into possible studies to corroborate this research, adaptations to the DSGS v2 model, and refinements to a computer implementation
- Summary: a review of the research project and findings

9.1 The validity of the theoretical basis

Four findings were drawn from the research literature (see p.39). These relate to: the contents of a training scheme designed to raise awareness of group skills, the initial ability of participants, issues for evaluation, and the appropriate medium for the training scheme. These were incorporated into the research project and validated by studies where possible. This section considers their accuracy and impact.

9.1.1 Conclusion 1: "Training schemes that focus on communication, responsibility and reflection will result in improved cognitive and social skills"

Gallimore and Tharp (1990) state that for a child to transfer skills between group activities he must be able to achieve the group's goal *and* be able to adapt the script from which he is working. Group activities in which this can occur are *joint activities-joint outcome* tasks. In these all the

children are required to: i) participate in order to complete the task, and ii) understand every aspect. The analysis of these activities in the literature review showed for a group to successfully complete such a task the individuals had to:

- *Communicate* – the group members exchange information, ask questions etc.
- *Take responsibility* – for their own work and that of the other group members
- *Reflect* – necessary for understanding and the internalisation of knowledge

The analysis of training schemes in Section 2.3 found that they supported at least one of these attributes. Furthermore, the children and the teachers participating in Study 1 mentioned that group work developed and required communication, responsibility and reflection – although the children from these schools did not employ the terms directly, nor did the teachers explicitly teach these skills.

On this basis, these training scheme principles were incorporated into the DSGS models. (Tables 4.1 and 5.1 summarise the methods to encourage communication, responsibility and reflection.) Evaluating the effectiveness of most of the existing schemes was precluded by the adaptations required and combination with other schemes. In the Aliens software based on the DSGS v2 model, for example, the children were not constantly exposed to the rules for effective group working (Dawes *et al.* 2000), and the procedural prompts were given with roles assigned (Palincsar *et al.* 1993, Goodman *et al.* 1998). The exception is a confirmation of the benefit of behaviour-based interactions, where the raising of awareness of appropriate behaviour before a group activity has been shown to improve task performance (Farivar and Webb 1994, Ashman and Gillies 1997, Wegerif and Dawes 1998). In this case the significantly better TGT performance in The Escape from the Forbidden Forest puzzle (Study 2) compared to the control group could be attributed to the initial consideration and discussion of appropriate behaviour. The impact of communication and reflection during the task and at the end is less as the explicit focus on the attributes appeared not to occur during the task - many children who used the training scheme failed to record group skill usage despite the prompts, and the assessment of overall group skill usage occurred after the puzzle had been completed.

This better TGT performance supports the first part of the conclusion that emphasising communication, responsibility and reflection, in this case in the GST *and* the TGT improves

individual cognitive skill. The conclusion is strengthened by the trend of improved recall in the Information and Communication Technology (ICT) quiz when only the GST had focused on these attributes (Study 6a). From a comparison of group activities in the classroom, a laboratory-setting and using a variety of media for recording the self-assessments, it was found the training scheme appeared not to impact on the group's TGT performance and the level of external support required (Studies 3a, 5, and 6b).

With the evaluation methods chosen, the second claim, that using a training scheme focusing on communication, reflection and responsibility would lead to improved social skills, was not substantiated. An analysis of the discourse and the TGT performance in group activities showed no difference between children who had used the training scheme for 18 weeks and a control group (Study 6b). This was corroborated by the teacher's opinion of the same children as she found a similar level of improvement in both groups (Study 6d).

In summary, the research findings provided partial support for the conclusion drawn from the literature review that training schemes which focus on the attributes of communication, responsibility and reflection result in improved cognitive and social skills. Children and teachers echo the importance of these attributes, and supporting them appeared to have:

- A positive impact on the individual cognitive ability
- A neutral impact on social skills and TGT performance in other group activities

9.1.2 Conclusion 2: "Children aged around 10 have sufficient metacognitive abilities to reflect on their behaviour with support and consequently amend their future actions"

As discussed in Section 2.3, existing training schemes designed to improve group skill usage frequently aim to develop metacognitive abilities. These schemes support the children in: reasoning about their actions and strategies, using resources, taking roles, monitoring and predicting outcomes, and understanding and explaining their actions (Wild 1996). Children require metacognitive abilities to adapt or reconceptualise the script for working together, which is needed to amend their future behaviour. Metacognition is necessary for appropriation. The child – the learner – must reason why the adult or computer – the expert – made a move or gave certain feedback to develop an interpretation of the expert's model (Newman *et al.* 1989).

Azmitia (1998) and Sharples (1999) suggest that metacognitive abilities emerge around the age of 10, as children are capable of negotiating roles and have the necessary vocabulary and skill to discuss their thinking. In the discussions in Studies 1b and 3b evidence of this linguistic ability is demonstrated by some children distinguishing between "giving help" and "giving the answers", and identifying cumulative and exploratory talk (Mercer 1994). In the presence of an adult, these children elaborated their ideas about groups and group work without prompts. This dialogue may have been as a consequence of the adult's presence but the resulting exchange of ideas by externalising their beliefs appeared to lead to the children internalising and formalising their understanding of groups (Pilkington and Jones 1996, Webb 1991, Chi *et al.* 1989).

This ability to use language explicitly was not shown by all of the 9 and 10 year olds who participated. In the WISP task and interviews (Studies 1a, 1b and 1e) the children appeared to understand why group work is useful; that is, they listed the benefits and gave rules for appropriate behaviour usually using language accepted by their culture (Wertsch and Stone 1985). In the WISP task this knowledge contrasted with their self-assessments (Study 1a). The children said they worked well together, indicating they believed they were listening and providing explanations. The observed behaviour, however, showed repeated interruptions and lack of elaborations. This indicated that without support, the children in the study lacked the ability to use metacognitive relating to group work appropriately and reason about their actions.

As a result of this observation, methods for supporting the children's use of language and their ability to reason about their actions were incorporated into the DSGS models. The Aliens software based on the DSGS v2 model *structured* the activity by requiring the children to consider group skill usage before, during and after the task (Tharp 1993). The prompted discussions should have improved the TGT performance; for example, the children acting on their plan to increase a group skill usage within the TGT. It *modelled* the desired type of exchanges in the discussion prompts to encourage reflection and communication (ibid). If the prompts are followed the children must: listen, take turns, provide justifications and check for understanding. It also provided *feedback* on group skill usage generated from the self-monitoring and the time the GST prompts were displayed (ibid). By adding support to raise the awareness of group skills it was hypothesised that the children would internalise the prompts and discuss definitions. This would lead to an interpretation of the expert model of the relationship between their behaviour and the generated feedback (Newman 1990, Vygotsky 1978).

Despite this support, the children did not seem to be able to switch between the GST and the TGT. The recordings and time the GST screens were displayed in classroom and laboratory-settings (Studies 4a, 4b and 5) indicated that the written prompts to focus the discussion, role assignment and modelling of appropriate exchanges were ineffective. As described in Section 7.3, many children were confused about what was being asked, while the majority appeared to focus on the TGT. This supports the findings of Brown (1987) and Schön (1987). Both reported that children have difficulty in reflecting on an activity while it is being performed.

The decrease in the average time spent on the overall self-assessment and the recorded discussions again suggests undeveloped metacognitive abilities (Study 5). The interviews and times the screens were displayed showed the majority of children believed that: i) the Aliens had been monitoring their behaviour directly, and ii) the computer-generated numeric score was accurate. These beliefs imply a naïve view of the infallibility of computers but also suggest a computer automatically has an expert status. Despite this, appropriation did not occur (Newman *et al.* 1989). Instead, the children treated the score as a summative mark and compared their results with each other.

However, there is evidence to suggest that the children's metacognitive abilities are developing within a classroom context. The numeric assessment of group skill usage made by the teacher observing the activity and that generated from the self-assessments (with account taken for overestimations) correlated. This is discussed in Section 7.4. This suggests children have begun to be able to reason about their behaviour. The encouragement to reflect on behaviour provided by the model may have led to the improvement in consistency between the self-assessments during, and at the end, of the group activity. Over the 18 weeks in which the scheme was used these two sets of self-assessments became more consistent for the majority of children, and the teachers noticed an improvement in their general group skills. This is discussed in Section 7.5.

In summary, the differences between the metacognitive abilities displayed in various group activities suggests that children aged about 10 are *beginning* to develop such skills. They need support to reflect on their behaviour. This conclusion acknowledges that behaviour is influenced by other factors, such as the presence of an adult, the perception of the computer, and the impact of written prompts to support children in reflecting on and rationalising about their behaviour. A greater improvement in group skill usage may have been achieved if children from year 6

participated, as proposed in Section 1.2.2, as at this age they appeared to have more metacognitive skills (see Study 1a, 1b and 1c) – although the limitations discussed in Section 8.1.4 still exist.

9.1.3 Conclusion 3: "The context will have an impact upon the success of any training scheme"

The idea that context influences behaviour is reinforced by the sociocultural and Activity Theory perspectives discussed in Sections 2.2 and 2.4. These emphasise the importance of the activity setting. Every activity outcome is unique as each setting is different; it depends upon the subject, the rules he follows, the community to which he belongs, the division of labour, the mediating artefact (including the delivery medium), his expectations of the object to be achieved and the interactions between these factors (Daniels 2001a, Engeström 1999a). Activity settings are not static, these factors and their relationships will change over time (Engeström 1999c).

The view that the interaction between factors affects the outcome was supported throughout the research project. The observations and answers given by the child show he is strongly influenced by the classroom culture, i.e., the hidden curriculum (Branco 2001), which is created by a teacher – the expert on classroom practice. Examples of this influence include the:

- Similar group formations and subjects chosen by pupils and their teacher (Study 1)
- The classification of cumulative talk as benefiting group work more than exploratory talk in a school where quiet working was encouraged (Study 3b)
- The pupils' focus on the TGT for the majority of time (Studies 2, 3a, 5 and 6)

The schools that volunteered to participate in this project all had teachers who used a didactic teaching style. From the above discussion it can be concluded that this would have influenced any group activity. It is hypothesised that in other environments the scheme would have a different impact. A collaborative school as described by Daniels (2001b), for example, would place more emphasis on the development of group skills, therefore implying the children would focus more on the GST.

The outcome is also affected by the delivery medium. This was illustrated by a comparison of self-assessments by the same children using the same training scheme and similar classroom-

based TGTs, but recording their self-assessments on a computer one week and on paper the next. It was found that when using a computer the children were significantly more likely to give higher self-assessments during the task, and have a significantly lower opinion of group skill usage overall (Study 4b).

Determining whether repeated exposure influenced the impact of the training scheme raised another context-related issue: how the method used to evaluate the outcome influences the findings. An example of the possible impact is the evaluation of the Aliens software. The software was used in 45-minute long ICT lessons taught over 18 weeks, which were held in the computer laboratory and taught by the researcher. No impact on the development of group skills was observed in an analysis of the teacher's opinion and the TGT performance and discourse (Studies 6b and 6d). However, the log file examinations showed that for the majority of children the consistency between self-assessments made during and after the activity increased (Study 5).

It can be valuable to incorporate triangulation into research projects, as the use of multiple methods when gathering data minimises errors that may distort the conclusions (Cohen and Manion 1998, pp.233-251). Prior research on evaluating the impact of group activities has shown that findings are dependent upon when the evaluation occurred, what the activity was, and the medium used to assess performance (e.g., Howe 1991, Light and Littleton 1999d).

Taking account of the factors discussed above meant that care had to be taken within the intervention programme (Study 6) to ensure that the tasks used to assess group skill usage could be treated as joint activity-joint outcome tasks. This would allow the children to reflect on the script for working together and on the TGT. Furthermore, the tasks needed multiple methods of measurement that were easily quantifiable. Triangulation in this case was achieved through a comparison between the number of moves, the amount of help, the time and the discourse during the activity. The task itself should be seen as real, relevant and motivating, so that all group members would participate (Biott and Easen 1994c). Despite the care being exercised in choosing a task,the same activity can cause different outcomes. This was illustrated by the analysis of interactions and task performance of two girls with equal exposure to the training scheme (Study 6c). In one case the discourse became more exploratory and task performance worsened, while in the other discourse became less exploratory and task performance improved. Using an Activity Theory framework, which has as a key concept that the factors in an activity

setting interact, *and* general experimental considerations, these differences could be attributed to the pupil (the subject), his partner, his previous experience and his expectations etc. (Engeström 1999a, 1999c, Cohen and Manion 1998, p.170-172).

Finally, although not explicitly represented or mentioned within Activity Theory (Engeström 1999c) the location of the task appeared to influence the outcome of the activity. During the design phase of the Escape from the Forbidden Forest the children constantly monitored their behaviour and recorded their opinions. When the same activity was performed in a classroom such automatic recording was scarce (Study 2). The laboratory-style setting also appeared to influence the consistency of self-assessments when using pen and paper to record opinions (Studies 4a and 4b). The self-assessments about the amount of group work displayed by children in this setting were more likely to be consistent during and after the activity (see Figure 6.1). The same training scheme implemented with pen and paper used within a classroom did not generate this consistency (see Figure 6.3), and neither did using a computer to record opinions in the laboratory-style setting. These findings suggested other factors needed to be considered. One explanation is that direct observation by an adult influences the group dynamic (Biott and Easen 1994b, Pilkington and Jones 1996). The researcher's presence could have reminded the children that they were being monitored for group skill usage rather than TGT performance. However, the medium is equally influential, as shown by the similar values recorded on a computer.

In summary, the research project found the context strongly impacted upon the success of the training scheme in raising the children's awareness of group skills. That is the setting, in addition to the factors mentioned directly by Activity Theory of subject, mediating artefact, community, rules and division of labour, is influential. To measure the impact of a training scheme ecologically valid studies must be conducted, i.e., the scheme should be integrated within the classroom culture. The studies corroborate the claim that the evaluation method influences the outcome, which means that assessment tasks must be chosen carefully to reflect the skills that are to be tested. It also supported existing research that multiple assessment methods, triangulation, are required for a valid evaluation.

9.1.4 Conclusion 4: "Computers can act as a mediating artefact within a group setting"

Computers have been found to: encourage talk, structure activities to sustain a collaborative learning process, provide a physical reference that can be used in discussion, allow the group to

control the rate of the task, and motivate children (Light 1993, Kutnick 1994, Crook 1994, Wild and Braid 1996, McLoughlin and Oliver 1998, McFarlane 1997a, etc.). The idea that a computer was a suitable medium for group activities was reinforced by interviews with the participating teachers and pupils (Study 1). Both groups mentioned ICT as being appropriate for group activities. These reasons indicate that a computer can act as a mediating artefact within a group setting. However, the issue addressed by the research project is whether it is an appropriate medium for the training scheme.

From a technical stance implementing the DSGS model to support the GST on a computer is straightforward. The Aliens implementation of the DSGS v2 model contained a database of rules, generated reminders to self-assess, selected prompts and randomly assigned roles based on the children's inputs, logged the data (i.e., recorded the names, rules, self-assessments and timings), and calculated the feedback. Furthermore, the necessary attributes of communication, responsibility and reflection discussed in Section 9.1.1 can be incorporated. Computers encourage discussion, procedural prompts and roles can be generated to encourage reflection, and dual-key control enables each child to take responsibility for his self-assessment (Light 1993, Hoyles *et al.* 1991b, Wegerif 1996b, Palincsar *et al.* 1993, King *et al.* 1998, Stewart *et al.* 1999, etc.). The software can also use anthropomorphism, by having the Aliens act as observers and prompting discussion, which motivates children (e.g., White *et al.* 1999, Hoppe *et al.* 2000).

Despite the ease and advantages of a computer implementation the standard prompts in the Aliens software failed to generate discussion about differences in opinion. This failure contrasted with previous research, which found pausing a computer-based activity resulted in more discussion than a formal instruction to talk by the teacher (Enyedy *et al.* 1997, Wegerif 1996b). The absence of talk may be explained by the significantly fewer differences in self-assessments during the task when a computer was used to record opinions compared to paper (Study 4a). Using paper the children were significantly more likely to differ in opinion and spontaneously discuss their reasons. Even with procedural prompts and assigning roles the resulting discussions rarely led to the generation of explanations and improved understanding reported when using some CSCL systems (e.g., White and Frederiksen 1998, Lester *et al.* 1997, Anderson *et al.* 2000). Activity Theory argues that the medium is only one factor that could explain the decreasing time spent with the GST prompts displayed. Equally important is the subject's metacognitive skills,

these may prevent him switching between tasks, and his knowledge of the classroom context may mean he focuses on the TGT.

Another factor influencing the time spent responding to the GST prompts is related to the children's previous experience. The majority of computer-based activities observed in the participating schools produced summative assessment as in "drill and skill" packages, or success could be measured by the number of websites found (Study 1d). This experience could explain the agreement of opinions before entering self-assessments during the task – to ensure the "right" answer, or interpreting the computer-generated numeric feedback as a summative score, and the absence of prompted discussion.

The short display time of the GST prompts was improved in the Aliens implementation by using the children's preconceptions about computers. From interviews it seems that the children believed the computer was capable of observing them and gave additional prompts when they failed to answer the GST prompt (Study 5). The validity of this belief is supported by the increased time that these prompts were displayed. This increase and the inferred discussion may indicate a child views a computer in a similar fashion to an observer (Pilkington and Jones 1996); which may explain why, with only indirect access to the activity, the computer appears to be able to mediate the discussion (Howe and Tolmie 1998).

Even when discussion between the children occurs, a computer implementation of the training scheme is limited. A teacher is required despite the scheme being designed to: i) assist in the appropriation of feedback (discussed in Section 4.1.4), and ii) to enable self-regulation so the group is more autonomous (e.g., Bennett and Dunne 1992d). This is because the scheme cannot monitor the children's exchanges to ensure misconceptions are not internalised, nor arbitrate when required (Johnson *et al.* 1990, Crook 1994).

In summary, from a technological standpoint computers can act as a mediating artefact for groups as they can support the stages within the DSGS models. From an educational viewpoint further work is required, as a preconception appears to exist in children about the role of a computer and the feedback provided. Computers can influence, but not monitor, behaviour – a teacher is still required to ensure group skills are understood.

9.2 The research project contributions

In this section the contributions towards the fields of understanding group work and CSCL are discussed. These educational, technical, and methodological implications are separated from the study outcomes as they form the principles on which future work could be based once they have been validated.

The first contribution is an educational finding relevant to the domain of evaluating group work. It was found that:

- The self-assessments made of group skill usage by 9 and 10 year olds are sufficient to generate a numeric value that correlates to the opinion of a teacher observing the same group activity
- As the children's self-assessments during the activity become more consistent with those after it, the correlation between the numeric value generated by the system and the teacher's opinion improves

These findings imply that the training scheme could *corroborate existing evaluations of group skill usage* as required by the National Curriculum (DfES 2000). This method of evaluation is beneficial as:

- Gathering the information required does not interfere with the TGT
- The calculation requires minimal data and does not involve a complex analysis
- It can be calculated for any group activity in a domain

The second educational contribution is that a training scheme based on the DSGS v2 model appeared to reflect group skill usage. This is shown by a comparison with the classification of changes in group skill usage by the teachers between the start and end of the study. The majority of children who became more consistent in self-assessments were said to have improved in group skill usage, while the majority of those who become less consistent stayed the same in terms of group skill usage. This suggests that the training scheme could be used *as a diagnostic tool.*

This relationship between the self-assessments and group skill usage has technical implications for CSCL systems. The use of computers to record the self-assessments and the time the screens were displayed provides *a new approach to evaluating group skills in CSCL.* In this approach the children are responsible for identifying group skills and the criteria for their assessment, and with

prompts and clearly defined assessment scales to monitor their behaviour during and at the end of the task. It is hypothesised that these self-assessments will take into account factors to which the computer has no direct access, such as the elaborate verbal responses to a prompt (Enyedy *et al.* 1997), and information that is expressed non-verbally (Cummings 1988). This information, with the time the children take to reflect and the assumed discussion times recorded by the computer, is used to generate a numeric feedback value and discussion prompt. This approach to generating feedback appears easier for children than systems that analyse the typed responses (Soller 2001, Soller *et al.* 2002, Baker and Lund 1997). Children using such systems were found to have difficulty expressing their opinion, even when using pre-defined sentence openers (Robertson *et al.* 1998).

The fourth contribution identifies the limitations of a computer-based training scheme to raise children's awareness of their skills at interacting face-to-face. Implementing the training scheme on a computer has many advantages: a computer is seen as a motivating environment suitable for groups, it could support communication, responsibility and reflection, and use the input to efficiently generate prompts and feedback. From the input times, additional prompts could be generated when the discussions appear to be short, giving the impression of being observed and possibly encouraging discussion and reflection (Pilkington and Jones 1996). Additionally, the data gathered could be used to indicate an absence of knowledge of group skills, illustrated by a child constantly awarding himself the maximum self-assessments and viewing the GST screens briefly. Despite this, a computer cannot monitor the children's behaviour; it can only raise an awareness about the interactions. Any computer-based training scheme in this situation requires that *the teacher monitors the groups and when necessary reminds and assists the children to develop definitions of group skills and reflect upon their behaviour.*

The final contribution is a methodological one. It relates to the sociocultural and Activity Theory perspective that underpins this research. The studies confirmed that changing one factor when using the training scheme impacted upon the outcome. This emphasises that the *educational impact of using the training scheme in a classroom can only be assessed through an ecologically valid study* (Kutnick and Baines 2001). In such studies it is necessary to have multiple measurements to evaluate group skill usage, as the method of assessment impacts the findings.

9.3 Future work

Suggestions for future work can be divided into three areas. In Section 9.3.1 further studies using the existing model are discussed. They address the issues of reliability and impact of the context on the outcomes. Section 9.3.2 considers changes to the model that arose from the studies. This includes the impact of written prompts and attitude towards, and calculation of, the numeric feedback. In the final section, 9.3.3, work relating to a computer implementation, that is, the appearance of the software, the provision of data, and the possible integration with distributed CSCL systems are considered.

9.3.1 Corroboration of studies

The areas for future work proposed in this section are further empirical studies to validate the existing findings. Currently caution is needed when extrapolating these findings as the quasi-experimental approach limited the number and type of schools participating. Technical difficulties and study mortality further reduced the number of participants (see the discussion of limitations in Section 8.1.4).

1. The validity of the heuristic

The small number of participants and the subjectivity when measuring group skill usage suggests that Study 5 should be repeated to corroborate the heuristic (Cohen and Manion 1998, p.71). This is necessary as accuracy is required for the appropriation process. The study would require several teachers to assess the groups simultaneously to confirm the reliability of their opinions, as these were found to differ between schools (p.88). These additional teacher evaluations of group skill usage and children's self-assessments could confirm the reported findings that:

- The numeric feedback generated correlates to the teacher's opinion of group skill usage
- A relationship exists between observed group skill usage and the self-assessment evaluations
- The training scheme can be used as a diagnostic tool

2. The validity of the findings

The second proposed area addresses the problems in extrapolating the findings. Currently these can only be applied to children aged 9 and 10, who are taught didactically, and where the focus is on domain tasks. In Section 9.1.4 it was hypothesised a different culture would impact the focus on the GST (Daniels 2001b). Repeating Studies 5 and 6 with older children and from a wider

range of classroom cultures would verify this. These repeated studies could indicate the impact of age and the environment on the:

- Consistency and scores for the self-assessments
- Amount of time spent on the GST and the TGT
- Children's beliefs about the validity and purpose of the feedback

3. Incorporation into classroom practice

A third limitation of the research project was that it did not employ an Action Research approach. The teachers did not act as co-researchers, contribute to the development of the model, and integrate the training scheme into their existing teaching practices. In effective schemes for raising an awareness of group skills the teacher uses group work for a variety of subjects (e.g., Slavin 1996). To perform a comparison with existing schemes to improve group skill usage a study would needed in which the teachers would display the rules for effective group work in their classrooms, model appropriate behaviour, and ensure the children performed the GST. This incorporates design principles from existing research, which stressed that training schemes only benefit students if used repeatedly and in multiple domains (e.g., Dawes *et al.* 2000, Adey and Shayer 1994). It is hypothesised in this environment the training scheme would lead to improved cognitive performance *and* develop social skills.

4. Confirmation of linguistic and metacognitive abilities

The final research area relating to the existing training scheme arose from Section 9.1.1. It is necessary to determine the accuracy of the children's self-assessments. The WISP task (Study 1a) suggested these are due to children not noticing incongruities in their behaviour (Johnson-Laird 1986), or using language to describe group skills differently to more experienced members of the culture (Wertsch and Stone 1985). However, when the task is to categorise scripts displaying disputational, cumulative and exploratory talk these children successfully evaluated the scripts according to the amount of group skills displayed (Mercer 1994) (Study 3b). This demonstrated that: i) when the TGT incorporates a GST task, and support in the form of prompts are provided, children are capable of the reflection and communication required by the GST and, ii) that although simple, such categorisation is robust enough to be used by children without training.

A further study is required to determine if children are equally accurate in assessing complex scenarios as the TGT. To make it a realistic task it is proposed that children are recorded in a lesson where they use the Aliens software. Next they would be asked to watch the video and perform the same self-assessments with questioning by an adult to identify their reasoning. The aim of this activity is to determine if the inaccurate self-assessments are attributable to difficulties in: switching between the TGT and GST, recognising incongruities between their words and behaviour, or understanding the concept of group skills. In this latter case it would be necessary to have the class teacher assess the groups using the same prompts. She strongly influences the child's understanding, and the child is likely to have internalised her concept of group work. A comparison of the children's assessments and teacher's opinion of different groups would indicate whether children are more objective when considering others (Dillenbourg 1999).

9.3.2 Adaptations to the DSGS v2 model

In this section issues are addressed that arose from Study 5 relating to: the time spent on the GST, the appropriateness of the written prompts, the attitude towards the feedback, and the contents and weighting within the heuristic.

1. Increasing the time spent on the GST

From the 18 weeks of data gathered there was a slight trend for the length of time spent on Stage 1 (choosing appropriate rules for the group activity and criteria for assessment), and Stage 4 (deciding final self-assessment score) to decrease. A greater trend was observed in the decreasing time to reflect on an appropriate self-assessment in Stage 2. There was no significant change in the average discussion of scores time in Stage 2. (The times taken are summarised in Appendix K.) The reduction in time implies that the children are reflecting and communicating less. These attributes have been found to be necessary to raise an awareness of behaviour. The reduction could be attributed to the children being more aware of what was required, not reading the instructions, referring to previous examples, and using existing agreed criteria (as discussed in Section 7.3.2). Possible adaptations to the software to increase the time spent on Stage 1include:

- Varying the prompts, e.g., randomly select an introductory screen so that the children cannot assume the content and must read the instructions
- Providing alternative introductions to the task. This may satisfy the requirement to ensure the task is fully understood (Lin 2001, Kramarski and Mevarech 1997, Salomon 1984)

- Extending the database and varying the rules displayed so that the children cannot select the same rules each week and state they are using previously agreed criteria
- Generating additional prompts, for example, asking the children to explain the assessment criteria if little time is spent selecting the rules. In children aged 9 and 10 it is hypothesised that this additional prompt could be interpreted as the computer monitoring them, which leads to more detailed self-explanations and discussions (Pilkington and Jones 1996)
- Requiring the children to decide and input rules for working together. This violates the design principles of minimal typing and turn taking (Wegerif and Dawes 1998, Druin 2002) but encourages elaboration and understanding through generating a rule and assessment criteria (Chi *et al.* 1989, Ericsson and Simon 1998, Webb 1991, Study 1e, Study 3b)

The prompts during the activity in Stage 2 could also be refined to increase the time spent on the GST and prevent the statements of: "But Miss… we've done that". Possible suggestions to investigate include:

- Varying self-assessment prompts: instead of asking the proportion of time the children listened to each other, they could be asked how often they interrupted their partner in the last five minutes. This requires adapting the scale, but these could be mapped to the existing 5-point scale used by the heuristic. This would ensure the screen is read *and* would demonstrate the type of criteria they are expected to use (Ager 1998)
- Adapting prompts to take into account previous responses. The children could discuss why their opinions changed over time, as well as why they differed
- Giving roles to be performed within the TGT if the children give a low mark instead of asking for a plan to improve in future. That is, the prompt could state one partner should always check that they agree before giving an answer, or that they must take it in turn to say why they are performing an action (Burton *et al.* 1997, Goodman *et al.* 1998, etc.)

2. Presenting the numeric feedback

The participating children appeared not to respond to the prompt designed to assist in appropriation, i.e., they did not use the feedback to develop a framework of understanding similar to an expert to represent the relationship between behaviour and opinion of group skill usage (Newman *et al.* 1989). Instead the children compared the computer-generated numeric feedback with others in the class; their aim was to have a higher score, or a more similar opinion

between their score and the numeric feedback. Preconceptions of computer feedback discussed in Section 9.1.4 indicate that as it was numeric the feedback was viewed as a summative mark. However, a numeric representation was effective as the children recognised that a higher score meant more time was spent using group skills (Gallimore and Tharp 1990). One possibility is to change the number to a written comment that is equally easy to interpret. A score of 10% from the Aliens could become: "You did not work together at all well today". These prompt questions could be extended if the system kept a record of responses. The marks could be contrasted between weeks or a child could be asked to explain why he works better with one partner rather than another.

3. Contents and weighting within the heuristic

The final area relating to the DSGS v2 model refinements was mentioned in Section 7.4.2. In this section it was suggested amending the bonus marks within the heuristic to reflect the actual times the GST screens were open. The accuracy of the feedback is important. It was hypothesised that by raising the child's awareness of their misconceptions by giving an "expert" opinion the children would discuss and appropriate a refined understanding of group skills. As their consistency increases it becomes more important that this value accurately reflects the observation of a teacher. Suggestions to improve the heuristic include:

- Incorporating the time to choose the rules: agreeing rules and criteria before the task (behaviour-based interactions) appeared to benefit task performance (Section 9.1.1). A longer time spent on this activity should imply a greater use of group skills
- Changing the weighting: currently those recorded during the task are given equal weight to the overall opinion, despite the children appearing to spend little time reflecting on their behaviour during the activity (Section 9.1.2). This suggests that more weight should be given to the final opinion, a proposal that is supported by existing research suggesting that it is the review of behaviour at the end of the activity that strongly influences future behaviour (e.g., Tann 1990, Johnson *et al.* 1990, Galton and Williamson 1994)

Sufficient material exists in the log files and teacher assessments for initial investigations into these suggested areas to be performed without conducting further studies.

9.3.3 Refinements to a computer implementation of the training scheme

The third area for future work addresses the impact and possibilities of the training scheme as a CSCL application. These are: the interface, enabling access to the log files, and integrating the training scheme into existing applications that assess group skill usage. It assumes that the reliability of the children's self-assessments and the relationship between these and the teacher's opinion has been confirmed.

1. The interface

Without changing the underlying model two adaptations to the interface are suggested for further study.

- Enabling the group to choose the animated character. Although the Aliens were popular with children of both sexes in Year 5, motivation may be improved if younger children can choose a character (Hoppe *et al.* 2000). The addition of choice may be important in implementations for older children, where animated characters might be beneficial (White *et al.* 1999) but smiling aliens seem childish (Lester *et al.* 1997)
- Displaying the rules chosen for working together. This is advocated in a classroom environment (Dawes *et al.* 2000) and when integrated into a computer-based group activity there was a significant impact in individual task performance (Study 2). When the children cannot view these rules the improvement in task recall is less significant (Study 6a). The sociocultural and Activity Theory approach indicates displaying the rules may be a contributing factor. Work is needed to determine their impact if displayed in a computer-based TGT. Do they benefit the performance or interfere with the TGT by restricting screen space?

2. Enabling access to the log files

Change in behaviour occurs over time (Adey and Shayer 1994, Engeström 1999c). The DSGS model design meant the outcome of group skill usage from one group activity would be considered when performing the GST in the next group activity (see Figures 4.2 and 5.1). In a computer implementation log files were created containing group information, self-assessments and timings. These log files provide a representation of the child's understanding, although to interpret the information, additional knowledge, such as the ability of the child, is required.

Enabling access to the log files may provide the teacher with information she could use to identify children who need extra support in generating criteria for measuring the usage of their chosen group skills and in relating their actions to these conditions. These children could be identified by: inconsistent self-assessments, constant perfect assessments and high overall scores (children with improved group skill usage often have low overall assessments – as shown in Section 7.5.1), spending little time viewing the GST screens, and creating possibly inappropriate rules for group work. A comparison over time would allow the teacher to identify children whose overall opinion repeatedly does not relate to the assessments during the task and therefore act as a diagnostic tool (Section 9.2). Finally, this information could be used to suggest possible future pairings of the children. Webb (1985) found that mixing high and medium, or medium and low achieving children resulted in more elaborated justifications. Again it should be noted that knowledge and understanding of group skills *and* the task *and* other group members influence behaviour (Daniels 2001a, Engeström 1999a). Teachers need to be involved in the development of appropriate representations of this collated data.

Any child could also interrogate the data stored about his group skill usage. He could look at the frequency of rules chosen, the change in self-assessments during the task and overall, the performance with different partners and the time spent with the GST screens displayed. The research project findings suggest scaffolding would be necessary to assist the children interpret this data. A record of the numeric feedback and opinions of group skill usage alone does not prompt reflection, as demonstrated by the competitive comparison of the paper record sheets (Study 5). The scheme could also give a limited amount of advice: if a child consistently records that he is poor at listening, the training scheme could suggest that they do not interrupt their partner.

A separate investigation is needed to assess the claim that it is beneficial to allow learners to access the data stored about their peers (presented anonymously) with the aim of monitoring their own progress (Kay 1997). Although beneficial for older learners, young children appear competitive about marks (Study 5) and are interested in determining who are in the anonymous groups (Robertson *et al.* 1998).

3. Integration into existing CSCL systems

The aim of this future work would be to support, rather than replace, current methods of evaluating participation or analysing exchanges in distributed CSCL systems (e.g., Constantino-Gonzalez and Suthers 2001a, 2001b, Soller 2001, Soller *et al.* 2002). It is proposed that it would triangulate with the existing analysis, and provide another method of explicitly encouraging an awareness of behaviour. According to the age, size and distributed nature of the group involved any implementation of the training scheme would have to consider:

- Adjusting the heuristic to take into account the extra self-assessments, as in distributed systems the group size is not constrained by the number of independent input devices as each learner has his own machine
- Re-evaluating the constants deducted to compensate for over-estimation, as if the group members are in secondary school or adults they should have appropriate vocabulary and the metacognitive abilities to reflect on their behaviour accurately
- Adapting the current rules suggested in Stage 1 so they are suitable for a distributed system
- Whether removing the discussion prompts is detrimental, as these may interfere with the TGT exchanges and as long as reflection occurred when performing the self-assessments the computer-generated numeric feedback appeared to correlate with the opinion of an observer
- How to display the analysis of the score and whether prompts or suggestions at this stage would benefit understanding

9.4 Summary

This research project involved the idea of separating a group activity into two tasks: one focused on the domain knowledge and the other on general learning skills. It was found that explicitly supporting the latter by separating the task into stages and specifically supporting the group to communicate, take responsibility and reflect appeared to lead to improved domain knowledge. This supports prior research findings that group work is beneficial to cognitive skills.

The initial claim that the research was novel as there is an interaction between the TGT and the GST is partially supported. Those children influenced by the feedback from the GST, demonstrated by an increase in the consistency of self-assessments during and after the TGT, were better at using group skills outside the training scheme in their teacher's opinion. The second claim that the approach was novel, that accurate feedback could be generated from self-

assessment, was supported. This feedback is necessary to ensure that the children's understanding of group skills and assessment of usage within the GST is reasonable. In previous research providing feedback required an analysis of the level of interactions, amount of help requested, discourse or final product.

To conclude, Activity Theory and sociocultural theoretical perspectives assisted in the design and interpretation of the results of this research. This research highlights that the child's age, his exposure to the training scheme and his expectations of computers have an influence on the awareness of his interactions and group skills. To determine the impact of a training scheme it is important that it be tested *in situ* to account for these factors. It was also found that the teacher remains indispensable for mediating the child's appropriation of these skills and ensuring he is able to transfer them to other environments.

REFERENCES

Adey P and Shayer M (1994) *Really raising standards: Cognitive intervention and academic achievement,* Routledge.

Ager R (1998) "Assessment, reporting and recording". In *Information and Communication Technology in Primary Schools: Children or Computers in Control?*, David Fulton Publisher, pp.47-63.

Anderson A, Cheyne W, Foot H, Howe C, Low J and Tolmie A (2000) "Computer support for peer-based methodology tutorials", *Journal of Computer Assisted Learning,* Vol. 16, No. 1, pp.41-53.

Anderson A, McAteer E, Tolmie A and Demissie A (1999) "The effect of software on the quality of talk", *Journal of Computer Assisted Learning,* Vol. 15, No. 1, pp.28-40.

Anderson A, Tolmie A, McAteer E and Demissie A (1993) "Software Style and Interaction around the Microcomputer", *Computers and Education,* Vol. 20, No. 3, pp.235-250.

Aronson E (2000) *Jigsaw Classroom,* <http://www.jigsaw.org/>.

Ashman A and Gillies R (1997) "Children's Cooperative Behavior and Interactions in Trained and Untrained Work Groups in Regular Classrooms", *Journal of School Psychology,* Vol. 35, No. 3, pp.261-279.

Azmitia M (1998) "Peer interactive minds: developmental, theoretical and methodological issues". In *Learning relationships in the classroom,* (Eds.) Faulkner D, Littleton K and Woodhead M, Routledge, pp.207-234.

Baker MJ and Lund K (1996) *Flexibly structuring the interaction in a CSCL environment*, Proceedings of the European Conference on Artificial Intelligence in Education, (Eds.) Brna P, Paiva A and Self J, pp.401-407.

Baker MJ and Lund K (1997) "Promoting reflective interactions in a computer-supported collaborative learning environment", *Journal of Computer Assisted Learning,* Vol. 13, No. 3, pp.175-193.

Barfurth M (1995) *Understanding the Collaborative Learning Process in a Technology Rich Environment: The Case of Children's Disagreements*, Proceedings of the Conference on Computer Support for Collaborative Learning (CSCL 95), (Eds.) Schnase J and Cunnius E, Lawrence Erlbaum Associates Ltd. pp.8-13.

BECTa (British Educational Communications and Technology agency) and NAACE (National Association of Advisors for Computers in Education) (1997) *Implementing IT Resource Pack,* BECTa.

Bellamy RKE (1996) "Designing Educational Technology: Computer-Mediated Change". In *Context and Consciousness: Activity Theory and Human-Computer Interaction,* (Ed.) Nardi BA, The MIT Press, pp.123-146.

Bennett N (1994) "Co-operative Learning". In *Classrooms in Groups in Schools,* (Eds.) Kutnick P and Rogers C, Cassell Education, pp.50-65.

Bennett N and Dunne E (1992a) "Introduction". In *Managing Classroom Groups,* Simon & Schuster Education, pp.1-17.

Bennett N and Dunne E (1992b) "Co-operative grouping". In *Managing Classroom Groups,* Simon & Schuster Education, pp.18-34.

Bennett N and Dunne E (1992c) "Group and classroom processes". In *Managing Classroom Groups,* Simon & Schuster Education, pp.35-63.

Bennett N and Dunne E (1992d) "Designing task: Cognitive aspects". In *Managing Classroom Groups,* Simon & Schuster Education, pp.64-90.

Bennett N and Dunne E (1992e) "Training in groupwork skills". In *Managing Classroom Groups,* Simon & Schuster Education, pp.134-170.

Biott C and Easen P (1994a) "Children Learning to be Together in Classrooms". In *Collaborative Learning in Staffrooms & Classroom,* David Fulton Publishers: London, pp.7-34.
Biott C and Easen P (1994b) "Children Learning to be Contributing Members of Classroom Groups". In *Collaborative Learning in Staffrooms & Classroom,* David Fulton Publishers: London, pp.35-66.
Biott C and Easen P (1994c) "Children Working in Groups to Solve Real Problems". In *Collaborative Learning in Staffrooms & Classroom,* David Fulton Publishers: London, pp.162-168.
Biott C and Easen P (1994d) "Understanding and Promoting Collaborative Learning". In *Collaborative Learning in Staffrooms & Classroom,* David Fulton Publishers: London, pp.203-209.
Boehm BW (1988) "A spiral model of software development and enhancement", *Computer,* Vol. 21, No. 5, pp.61-72.
Boulay B, Luckin R and Soldato T (1999) *The Plausability Problem: Human Teaching Tactics in the 'Hands' of a Machine*, Proceedings of the Conference on Artificial Intelligence in Education, (Eds.) Lajoie S and Vivet M, IOS Press, pp.225-232.
Branco AU (2001) "Contextual, Interactional and Subjective Dimensions of Cooperation and Competition: A Co-Constructivist Analysis". In *The Theory and Practice of Cultural-Historical Psychology,* (Ed.) Chaiklin S, Aarhus University Press, pp.107-123.
Brandes A and Wilensky V (1980) "Treasureworld: A Computer Environment for the Study and Exploration of Feedback". In *Constructionism,* (Eds.) Harel I and Papert S, Norwood MA: Ablex Publishing, pp.391-416.
The British Psychological Society (2000) *Code of Conduct, Ethical Principles & Guidelines,* <http://www.bps.org.uk/documents/Code.pdf>.
Brown A (1987) "Metacognition, executive control, self-regulation and other more mysterious mechanisms". In *Metacognition, Motivation and Understanding,* (Eds.) Weinert FE and Kluwe RH, Lawrence Erlbaum Associates, pp.65-116.
Brown AL and Campione JC (1987) "Linking Dynamic Assessment with School Achievement". In *Dynamic assessment: an interactional approach to evaluating learning potential,* (Ed.) Lidz CS, New York; London: Guilford Press, pp.82-115.
Brown J and Howlett F (1994) *NCET information technology worker: stimulate to educate,* National Council for Educational Technology.
Burton M, Brna P and Treasure-Jones T (1997) *Splitting the Collaborative Atom: How to Support Learning about Collaboration*, Artificial Intelligence in Education: Knowledge and Media in Learning Systems, (Eds.) du Boulay B and Mizoguchi R, IOS, Amsterdam, pp.135-142.
Chan T-W and Baskin AB (1990) "Learning Companion Systems". In *Intelligent Tutoring Systems: At the crossroads of Artificial Intelligence and Education,* (Eds.) Frasson C and Gauther G, Norwood Abbey, pp.6-33.
Chi MT, Bassok M, Lewis MW, Reimann P and Glaser R (1989) "Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems", *Cognitive Science,* Vol. 13, pp.145-182.
Cohen L and Manion L (1998) *Research Methods in Education - fourth edition,* Routledge.
Constantino-Gonzalez M and Suthers DD (2001a) *Coaching collaboration by comparing solutions and tracking participating*, Proceedings of the European Conference on Computer Supported Collaborative Learning, (Eds.) Dillenbourg P, Eurelings A and Hakkarainen K, Unigraphic Maastricht, pp.173-180.
Constantino-Gonzalez M and Suthers DD (2001b) *Designing and Evaluating a Collaboration Coach: Knowledge and Reasoning*, Proceedings of the Conference on Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future, (Eds.) Moore JD, Redfield CL and John WL, IOS Press, pp.176-187.

Cooper B and Brna P (2001) *Good relationships can lead to a long and happy engagement: supporting high quality interaction and motivation in the classroom ICT - the NIMIS project*, British Educational Research Association Annual Conference.
Crook C (1991) "An introduction to measuring and understanding the learning process", *Computers & Education,* Vol. 17, No. 1, pp.81-91.
Crook C (1994) *Computers and the Collaborative Experience of Learning,* Routledge.
Crook C (1998) "Children as computer users: The case of collaborative learning", *Computers & Education,* Vol. 30, No. 3/4, pp.237-247.
Crook C (1999) "Computers in the community of classrooms". In *Learning with computers: Analysing productive interactions,* (Eds.) Littleton K and Light P, Routledge, pp.102-117.
Cummings R (1988) "Small-group discussions and the microcomputer". In *Computers in Education 5 - 13,* (Eds.) Jones A and Scrimshaw P, Open University Press, pp.149-158.
Daniels H (2001a) "Vygotskian theory and education". In *Vygotsky and Pedagogy,* Routledge Falmer, pp.30-68.
Daniels H (2001b) "The institutional level of regulation and analysis". In *Vygotsky and Pedagogy*, Routledge Falmer, pp.131-175.
Daniels H (2002) *Personal Communication.*
Dawes L, Mercer N and Wegerif R (2000) *Thinking together: A programme of activities for developing thinking skills at KS2,* Questions Publishing Co Ltd.
DfEE (Department for Education and Employment) (1998) *Information and Communications Technology (ICT) in Schools,* <http://www.dfee.gov.uk/ict/index.htm>.
DfEE (Department for Education and Employment) Government paper (1999) *WEIGHING THE BABY: The report of the Independent Scrutiny Panel on the 1999 Key Stage 2 National Curriculum tests in English and mathematics,* <http://www.dfee.gov.uk/panel/>.
DfES (Department for Education and Skills) (2000) *The National Curriculum,* <http://www.nc.uk.net/home.html>.
DfES - (Department for Education and Skills) (2001) Schemes of work: Primary Information Technology, Department for Education and Skills <http://www.standards.dfee.gov.uk/schemes/it>.
Dillenbourg P (1999) "What do you mean by collaborative learning?". In *Collaborative learning: Cognitive and Computational Approaches,* (Eds.) Dillenbourg P, Oxford: Elsevier, pp.1-19.
Dillenbourg P, Baker M, Blaye A and O'Malley C (1996) "The Evolution of Research on Collaborative Learning". In *Learning in Humans and Machines: Towards and Interdisciplinary Science,* (Eds.) Reimann P and Spada H, Pergamon, pp.189-211.
Dillenbourg P and Self J (1995) "Designing Human-Computer Collaborative Learning". In *Computer Supported Collaborative Learning,* (Eds.) O'Malley C, Springer-Verlag, pp.245-264.
Doise W (1990) "The development of individual competencies through social interaction". In *Children helping children,* (Eds.) Foot H, Morgan M and Shute R, Chichester, J. Wiley and Sons, pp.43-64.
Draper SW and Anderson A (1991) "The significance of dialogue in learning and observing learning", *Computers and Education,* Vol. 17, No. 1, pp.93-107.
Druin A (2002) "The Role of Children in the Design of New Technology", *Behaviour and Information Technology,* Vol. 21, No. 1, pp.1-25.
Dunne E and Bennett N (1996) *Talking and Learning in Groups,* Routledge.
Edwards AD and Westgate DPG (1994a) *Investigating Classroom Talk (Revised and Extended Second Edition),* The Falmer Press.
Edwards AD and Westgate DPG (1994b) "Interpreting Classroom Communication: Turns, Sequences and Meanings". In *Investigating Classroom Talk (Revised and Extended Second Edition),* The Falmer Press, pp.101-133.

Edwards D and Mercer N (1987) *Common Knowledge: The development of understanding in the classroom,* Methuen.
Elliott J (1976) "Preparing teachers for classroom accountability", *Education for Teaching,* Issue. 100, pp.49-71.
Engeström Y (1999a) *Changing practice through research: Changing research through practice (Keynote Address)*, Centre for Learning and Work Research 7th Annual International Conference on Post-compulsory Education and Training, pp.1-26.
Engeström Y (1999b) "Activity theory and individual and social transformation". In *Perspectives on Activity Theory,* (Eds.) Engeström Y, Meittinen R and Punamäki R-L, Cambridge University Press, pp.19-38.
Engeström Y (1999c) "Innovative learning in work teams: Analyzing cycles of knowledge creation in practice". In *Perspectives on Activity Theory,* (Eds.) Engeström Y, Meittinen R and Punamäki R-L, Cambridge University Press, pp.377-406.
Engeström Y, Brown K, Christopher L and Gregory J (1997) "Co-ordination, co-operation and communication in courts: expansive transitions in legal work". In *Mind, Culture and Activity,* (Eds.) Cole M, Engeström Y and Vasquez O, Cambridge University Press, pp.369-385.
Enyedy N, Vahey P and Gifford BR (1997) *Active and Supportive Computer-Mediated Resources for Student-to-Student Conversations*, Proceedings of the Second International Conference on Computer Support for Collaborative Learning, (Eds.) Hall R, Miyake N and Enyedy N, University of Toronto Press, pp.27-36.
Ericsson K and Simon H (1998) "How to Study Thinking in Everyday Life: Contrasting Think-Aloud Protocols with Descriptions of Explanations of Thinking", *Mind, Culture & Activity,* Vol. 5, No. 3, pp.178-186.
Farivar S and Webb N (1994) "Helping and Getting Help - Essential Skills for Effective Group Problem Solving", *Arithmetic Teacher,* Vol. 41, No. 9, pp.521-525.
Finlayson H and Cook D (1998) "The Value of Passive Software in Young Children's Collaborative Work". In *IT for Learning Enhancement,* (Ed.) Monteith M, Swets & Zeitlinger Publisher, pp.106-120.
Fisher E (1993) "The teacher's role". In *Language, Classrooms and Computers,* (Ed.) Scrimshaw P, Routledge, pp.57-74.
Forman E and Cazden C (1985) "Exploring Vygotskian perspectives in education: the cognitive value of peer interaction". In *Culture, Communication and Cognition: Vygotskian Perspectives,* (Ed.) Wertsch J, Cambridge University Press, pp.323-347.
Forman E and McPhail J (1993) "9. Vygotskian Perspective on Children's Collaborative Problem-Solving Activities". In *Contexts for Learning,* (Eds.) Forman E, Minick N and Stone C, Oxford University Press, pp.213-229.
Gallimore R and Tharp R (1990) "Teaching mind in society: Teaching, schooling, and literate discourse". In *Vygotsky and education: Instructional implications of sociohistorical psychology,* (Ed.) Moll L, Cambridge University Press, pp.175 - 205.
Galton M (1998) *Reliving the ORACLE experience: back to basics or back to the future?* Coventry: Centre for Research in Elementary and Primary Education, University of Warwick.
Galton M and Williamson J (1994) *Group work in the primary classroom,* Routledge.
Garvin JW, Butcher AC, Stefani LAJ, Tariq VN, Lewis MHR, Blumson NL, Govier RN and Hill JA (1995) "Group Projects for First-year University Students: an evaluation", *Assessment & Evaluation in Higher Education,* Vol. 20, No. 3, pp.273-287.
Gillies R and Ashman A (1998) "Behavior and Interactions of Children in Co-operative Groups in Lower and Middle Elementary Grades", *Journal of Educational Psychology,* Vol. 90, No. 4, pp.746-757.
Golding W (1997) *Lord of the Flies,* Faber and Faber.

Goodman B, Soller A, Linton F and Gaimari R (1998) "Encouraging Student Reflection and Articulation using a Learning Companion", *International Journal of Artificial Intelligence in Education,* Vol. 9, pp.237-255.
Grove J and Williams N (1998) "Explorations in Virtual History". In *IT for Learning Enhancement,* (Ed.) Monteith M, Swets & Zeitlinger Publisher, pp.171-183.
Hammersley M and Atkinson P (1995) *Ethnography: Principles in practice, Second edition,* Routledge.
Henry J (2001) *Warning to cool the test frenzy*. The Times Educational Supplement, November 2nd, p.8.
Hoppe U, Brna P, Paiva A and Tewissen F (2000) *Networked Interactive Media in Schools*, Esprit funded i3 Project on Experimental School Environments, Contract No: 29301.
Howe C and Tolmie A (1998) "Computer support for learning in collaborative contexts: prompted hypothesis testing in physics", *Computers & Education,* Vol. 30, No. 3/4, pp.223-235.
Howe CJ (1991) "Explanatory concepts in physics: towards a principled evaluation of teaching materials", *Computers and Education,* Vol. 17, No. 1, pp.73-80.
Hoyles C, Healy L and Pozzi S (1994) "Groupwork with Computers: An Overview of Findings", *Journal of Computer Assisted Learning,* Vol. 10, No. 4, pp.202-215.
Hoyles C, Sutherland R and Healy L (1991a) "14. Children talking in computer environments: New insights into the role of discussion in mathematical learning". In *Language in Mathematical Education: Research and Practice,* (Eds.) Durkin K and Shire B, Open University Press, pp.162-175.
Hoyles C, Healy L and Sutherland R (1991b) "Patterns of discussion between pupil pairs in computers and non-computer environments", *Journal of Computer Assisted Learning,* Vol. 10, pp.202-215.
Hoyles C, Sutherland R and Healy L (1993) "Children talking in computer environments: New insights into the role of discussion in mathematics learning". In *Language in Mathematical Education Research and Practice,* (Eds.) Durkin K and Shire B, Open University Press, pp.162-175.
Hutchins E (1995) *Cognition in the Wild,* MIT Press.
Inkpen K, Ho-Ching W, Kuederle O, Scott S and Shoemaker G (1999) *"This is Fun! We're All Best Friends and We're All Playing." Supporting Children's Synchronous Collaboration*, Proceedings of the Conference on Computer Supported Collaborative Learning (CSCL), (Eds.) Hoadley C and Roschelle J, Lawrence Erlbaum Associates, pp.252-259.
Issroff K (1999) "Time-based analysis of students studying the Periodic Table". In *Learning with computers: Analysing productive interaction,* (Eds.) Littleton K and Light P, Routledge, pp.46-61.
Jackson A, Kutnick P and Kington A (2001) "Principles and practical grouping for the use of drill and practice programs", *Journal of Computer Aided Learning,* Vol. 17, No. 2, pp.130-141.
Johnson DW, Johnson RT and Holubec EJ (1990) *Circles of Learning: Cooperation in the Classroom,* Interaction Book Company.
Johnson R (1995) "Two cheers for the reflective practitioner", *Journal of Further and Higher Education,* Vol. 19, No. 3, pp.74-83.
Johnson-Laird P (1986) "*An artist constructs a science",* Times Literary Supplement, August 15th**,** pp.879-80.
Joiner R, Messer D, Light P and Littleton K (1998) "The Effects of Gender, Expectations of Success and Social Comparison on Children's Performance on a Computer-based Task", *Educational Psychology,* Vol. 18, No. 3, pp.319-326.
Jones K and Haylock D (1985) "Developing children's understanding in mathematics", *Remedial Education,* Vol. 20, No. 1, pp.30-34.
Kagan S (1992) *Cooperative learning,* Kagan Cooperative Learning.

Kay J (1997) *Learner Know Thyself: Student Models to Give Learner Control and Responsibility*, Proceedings of the International Conference on Computers in Education, Association for the Advancement of Computing in Education (AACE), (Eds.) Halim Z, Ottomann T and Razak Z. pp. 17-24.

King A, Staffieri A and Adelgais A (1998) "Mutual Peer Tutoring: Effects of Structuring Tutorial Interaction to Scaffold Peer Learning", *Journal of Educational Psychology,* Vol. 90, No. 1, pp.134-152.

Klinger S (1999) "Coding categories to record student talk at a multimedia interface", *Journal of Computer Assisted Learning,* Vol. 15, No. 2, pp.109-117.

Knussen C, Tanner GR and Kibby MR (1991) "An approach to the evaluation of hypermedia", *Computers & Education,* Vol. 17, No. 1, pp.13-24.

Kolb D, Rubin M and McIntyre J (1984) "Learning and Problem Solving". In *Organisational Psychology: An Experimental Approach to Organisational Behaviour,* (Eds.) Kolb D, Rubin M and McIntyre J, Prentice Hall, pp.31-49.

Kozulin A (1986) "The concept of activity in Soviet psychology: Vygotsky, his disciples and critics", *American Psychologist,* Vol. 41, No. 3, pp.264-274.

Kramarski B and Mevarech Z (1997) "Cognitive-metacognitive training within a problem-solving based Logo environment", *British Journal of Educational Psychology,* Vol. 67, pp.425-445.

Kress G (2002) *Mulitmodality, Sign-making and Learning,* Centre for Sociocultural and Activity Theory Research (SAT) National Seminar, 23rd April 2002, University of Birmingham.

Kruger A (1993) "Peer collaboration: conflict, cooperation or both?", *Social Development,* Vol. 2, No. 3, pp.165-182.

Kutnick P (1994) "Use and Effectiveness of Groups in Classrooms: Towards a Pedagogy". In *Classrooms in Groups in Schools,* (Eds.) Kutnick P and Rogers C, Cassell Education, pp.13-33.

Kutnick P and Baines E (2001) *Towards a social pedagogy of group work*, International Conference on Communication, Problem-solving and Learning.

Kutnick P and Rogers C (1994) "Groups in classrooms". In *Groups in schools,* (Eds.) Kutnick P and Rogers C, Cassell Education, pp.1-12.

Kuutti K (1996) "Activity Theory as a Potential Framework for Human-Computer Interaction Research". In *Context and Consciousness: Activity Theory and Human-Computer Interaction,* (Ed.) Nardi BA, The MIT Press, pp.17-44.

Laurillard D (1993) "Analysing teaching media: Introduction". In *Rethinking university teaching: a framework for the effective use of educational technology,* London: Routledge, pp.97-105.

Lester J, Converse S, Kahler S, Barlow S, Stone B and Bhoga R (1997) *The Persona Effect: Affective Impact of Animated Pedagogical Agents*, CHI 97, (Ed.) Pemberton S, ACM Press, pp.359 - 366.

Light P (1993) "Collaborative learning with computers". In *Language, Classrooms and Computers,* (Ed.) Scrimshaw P, Routledge, pp.40-56.

Light P, Foot T, Colbourn C and McClelland I (1987) "Collaborative interactions at the microcomputer keyboard", *Educational Psychology,* Vol. 7, No. 1, pp.13-21.

Light P and Glachan M (1985) "Facilitation of individual problem solving through peer interaction", *Educational Psychology,* Vol. 5, pp.217-225.

Light P and Littleton K (1999a) "Preface". In *Social processes in children's learning,* (Eds.) Light P and Littleton K, Cambridge University Press, pp.xiii-xvii.

Light P and Littleton K (1999b) "Introduction: Getting IT together". In *Learning with computers: Analysing productive interactions,* (Eds.) Littleton K and Light P, Routledge, pp.2-9.

Light P and Littleton K (1999c) "Gender Agendas". In *Social processes in children's learning,* (Eds.) Light P and Littleton K, Cambridge University Press, pp.52-72.

Light P and Littleton K (1999d) "Social comparison and learning". In *Social processes in children's learning,* (Eds.) Light P and Littleton K, Cambridge University Press, pp.73-90.

Lin X (2001) "Designing metacognitive activities", *Educational Technology Research and Development,* Vol. 49, No. 2, pp.23-40.
Linehan C and McCarthy J (2001) "Reviewing the "Community of Practice" Metaphor: An Analysis of Control Relations in a Primary School Classroom", *Mind, Culture, and Activity,* Vol. 8, No. 2, pp.129-147.
Lipponen L (2002) *Exploring foundations for computer-supported collaborative learning*, Proceedings of the Conference on Computer Supported Collaborative Learning 2002, (Ed.) Stahl G, Hillsdale, NJ: Erlbaum. pp.72-81.
Littleton K (1999) "Productivity through interaction: an overview". In *Learning with computers: Analysing productive interactions,* (Eds.) Littleton K and Light P, Routledge, pp.179-194.
Littleton K, Light P, Joiner R, Messer D and Barnes P (1998) "Gender, Task Scenarios and Children's Computer-based Problem Solving", *Educational Psychology,* Vol. 18, No. 3, pp.327-340.
McFarlane A (1997a) "Where are we and how did we get here?". In *Information technology and authentic learning,* (Ed.) McFarlane A, Routledge, pp.1-12.
McFarlane A (1997b) "Thinking about writing?". In *Information technology and authentic learning,* (Ed.) McFarlane A, Routledge, pp.108-120.
McLoughlin C and Oliver R (1998) "Maximising the language and learning link in computer learning environments", *British Journal of Educational Technology,* Vol. 29, No. 2, pp.125-136.
Meittinen R (1999) "Transcending traditional school learning: Teachers' work and networks of learning". In *Perspectives on Activity Theory,* (Eds.) Engeström Y, Meittinen R and Punamäki R-L, Cambridge University Press, pp.325-344.
Mercer N (1994) "The quality of talk in children's joint activity at the computer", *Journal of Computer Assisted Learning,* Vol. 10, No. 1, pp.24-32.
Mercer N (1995a) "Ways of Talking". In *The Guided Construction of Knowledge: Talk Amongst Teachers and Learners,* Multilingual Matters Ltd, pp.9-20.
Mercer N (1995b) *The Guided Construction of Knowledge: Talk Amongst Teachers and Learners,* Multilingual Matters Ltd.
Mercer N and Fisher E (1992) "How do teachers help children to learn? An analysis of teachers' interventions in computer-based activities", *Learning and Instruction,* Vol. 2, No. 4, pp.339-355.
Mercer N and Wegerif R (1999) "Is 'exploratory talk' productive talk?". In *Learning with computers: Analysing productive interaction,* (Eds.) Littleton K and Light P, Routledge, pp.79-101.
Mercer N, Wegerif R and Dawes L (1999) "Children's Talk and the Development of Reasoning in the Classroom", *British Educational Research Journal,* Vol. 25, No. 1, pp.95-111.
Moreno R and Mayer RE (2000) "Engaging Students in Active Learning: The Case for Personalized Multimedia Messages", *Journal of Educational Psychology,* Vol. 92, No. 4, pp.724-733.
MRC - Medical Research Council (2000) *A framework for development and evaluation of RCTs for complex interventions to improve health,* <http://www.mrc.ac.uk/pdf-mrc_cpr.pdf>.
NCET - National Oracy Project and the National Council for Educational Technology (1990) *Talking IT Through: A discussion document.*
Newman D (1990) "Cognitive Change by Appropriation". In *Cognition, Computing and Cooperation,* (Eds.) Robertson S, Zachary U and Black J, Ablex, pp.84-94.
Newman D, Griffin P and Cole M (1989) *The construction zone: Working for cognitive change in school,* Cambridge University Press.
O'Shea T and Self J (1983) *Learning and teaching with computers: Artificial intelligence in education,* Englewood Cliffs, NJ: Prentice-Hall.

Palincsar A and Brown A (1988) "Teaching and Practicing Thinking Skills to Promote Comprehension in the Context of Group Problem Solving", *Remedial & Special Education (Rase),* Vol. 9, No. 1, pp.53-59.
Palincsar A, Brown A and Campione J (1993) "First-Grade Dialogues for Knowledge Acquisition and Use". In *Contexts for learning: sociocultural dynamics in children's development,* (Eds.) Forman E, Minick N and Stone C. Oxford, Oxford University Press, pp.43-57.
Pauli C and Reusser K (1997) *Supporting collaborative problem solving: supporting collaboration and supporting problem solving*, Swiss Workshop on Collaborative and Distributed systems, (Ed.) Dillenbourg P, <http://tecfa.unige.ch/tecfa/tecfa-activities/loz97contrib.html>. .
Phelps E and Damon W (1989) "Problem Solving with Equals: Peer Collaboration as a Context for Learning Mathematics and Spatial Concepts", *Journal of Educational Psychology,* Vol. 82, No. 4, pp.639-646.
Pierce K and Gholson B (1994) "Surface similarity and relational similarity in the development of analogical problem solving: Isomorphic and nonisomorphic transfer", *Developmental Psychology,* Vol. 30, No. 5, pp.724-737.
Pilkington R and Jones CP (1996) "Interacting with computer-based simulation: The role of dialogue", *Computers in Education,* Vol. 27, No. 1, pp. 1-14.
Pilkington R (1998) "Dialogue games in support of qualitative reasoning", *Journal of Computer Assisted Learning: Special Issue on Qualitative Reasoning,* Vol. 14, pp.308-320.
Pilkington R (2002) *Computer-Mediated Argumentation in Higher Education: Developing Discussion Skills via Roles*, Proceedings of the Third International Conference on Networked Learning 2002, (Eds.) Banks S, Goodyear P, Hodgson V and McConnell D, pp.481-489.
Pillay H (1998) "An Investigation of the Effect of Individual Cognitive Preferences on Learning through Computer-based Instruction", *Educational Psychology,* Vol. 18, No. 2, pp.171-182.
Puntambekar S (1995) *Investigating the Effect of a Computer Tool On Students' Metacognitive Processes* Cognitive science Research Report No. 381, University of Sussex.
Reid J, Forrestal P and Cook J (1989) *Small group learning in the classroom,* Chalkface Press.
Robertson J, Good J and Pain H (1998) "BetterBlether: The design and evaluation of a discussion tool for education", *International Journal of Artificial Intelligence in Education,* Vol. 9, pp.219-236.
Rogoff B (1990) *Apprenticeship in thinking: cognitive development in social context,* New York;Oxford: Oxford University Press.
Ryan R and Deci E (2000) "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions", *Contemporary Educational Psychology,* Vol. 25, No. 1, pp.54-67.
Säljö R (1998) "Thinking with and through artifacts: The role of psychological tools and physical artifacts in human learning and cognition". In *Learning relationships in the classroom,* (Eds.) Faulkner D, Littleton K and Woodhead M, Routledge, pp.54-66.
Salomon G (1984) "On ability development and far transfer: a response to Pea and Kurland", *New Ideas Psychology,* Vol. 2, No. 2, pp.169-174.
Salomon G (1995) *What Does the Design of Effective CSCL Require and How Do We Study Its Effects?*, Computer Supported Collaborative Learning, < http://www-cscl-95.indiana.edu/cscl95/outlook/62_salomon.html>.
Saxe G, Gearhart M, Note M and Paduano P (1993) "Peer interaction and the development of mathematical understandings: A new framework for research and educational practise". In *Charting the Agenda: Educational activity after Vygotsky,* (Ed.) Daniels H, Routledge, pp.107-144.
Schön D (1987) *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions,* San Francisco: Jossey-Bass.
Schön DA (1983) *The Reflective Practitioner. How professionals think in action,* London: Temple Smith.

Scott SD, Mandryk RL and Inkpen KM (2002) *Understanding Children's Interactions in Synchronous Shared Environments*, Proceedings of the Conference on Computer Supported Collaborative Learning 2002, (Ed.) Stahl G, Hillsdale, NJ: Erlbaum, pp.333-341.

Sharples M (1999) "Becoming a writer". In *How We Write: Writing as creative design,* Routledge, pp.13-36.

Sharples M, Corlett D and Westmancott O (2002) "The Design and Implementation of a Mobile Learning Resource", *Personal and Ubiquitous Computing,* Vol. 6, pp.220-234.

Sheingold K, Hawkins J and Char C (1984) ""I'm the thinkist you're the typist"; the interaction of technology and the social life of classrooms", *Journal of Social Issues,* Vol. 40, No. 3, pp.49-61.

Slavin R (1991) "Synthesis of research on cooperative learning", *Educational Leadership,* Vol. 1, pp.71-77.

Slavin R (1995) *Cooperative Learning: Theory, Research, and Practice. Second Edition,* Allyn & Bacon.

Slavin R (1996) "Research on Cooperative Learning and Achievement: What We Know, What We Need to Know", *Contemporary Educational Psychology,* Vol. 21, No. 1, pp.43-69.

Smagorinsky P (1998) "Thinking and Speech and Protocol Analysis", *Mind, Culture, and Activity,* Vol. 5, No. 3, pp.157-177.

Soller A, Wiebe J and Lesgold A (2002) *A Machine Learning Approach to Assessing Knowledge Sharing During Collaborative Learning Activities*, Computer Supported Collaborative Learning Conference 2002, (Ed.) Stahl G, Hillsdale, NJ: Erlbaum, pp.128-137.

Soller A (2001) "Supporting Social Interaction in an Intelligent Collaborative Learning System", *International Journal of Artificial Intelligence in Education,* Vol. 12, No. 1, pp.40-62.

Spavold J (1989) "Children and Databases: An Analysis of Data Entry and Query Formulation", *Journal of Computer Assisted Learning,* Vol. 5, No. 3, pp.145-160.

Stephenson P (1997) "Children's learning using control information technology". In *Information technology and authentic learning,* (Ed.) McFarlane A, Routledge, pp.38-51.

Stevens R and Slavin R (1995) "The Co-operative Elementary School: Effects on Students' Achievement, Attitudes and Social Relations", *American Education Research Journal,* Vol. 32, No. 2, pp.321-351.

Stewart J, Bederson B and Druin A (1999) *Single Display Groupware: A model for Co-present Collaboration*, Proceedings of the Computer Human Interaction Conference (CHI 97), ACM Press, pp.286-293.

Stone CA (1993) "What is missing in the metaphor of scaffolding?". In *Contexts for learning: sociocultural dynamics in children's development,* (Eds.) Forman E, Minick N and Stone C. Oxford, Oxford University Press, pp.169-183.

Straus S (1997) "Technology, Group Process, and Group Outcomes: Testing the Connections in Computer-Mediated and Face-to-Face Groups", *Human Computer Interaction,* Vol. 12, pp.227-266.

Tann S (1990) "Awareness of Language". In *Teacher Handbooks: Talking and Listening,* (Eds.) Wray D, Wade B, Tann S, Phillips T, Armitage M, Carre C and Duffin J, Scholastic Publications Ltd., pp.137-159.

Tharp R (1993) "Institutional and Social Context of Educational Practice and Reform". In *Contexts for Learning,* (Eds.) Forman E, Minick N and Stone CA, Oxford University Press, pp.269-282.

Tomlinson P (1989) "Having it Both Ways: hierarchical focusing as research interview method", *British Educational Research Journal,* Vol. 15, No. 2, pp.155-176.

Tudge J (1990) "Vygotsky, the Zone of Proximal Development and Peer Collaboration: Implications for classroom practice". In *Vygotsky and Education,* (Ed.) Moll LC, Cambridge University Press, pp.155-172.

Underwood G, Underwood J, Pheasey K and Gilmore D (1996) "Collaboration and discourse while programming the KidSim microworld simulation", *Computers and Education,* Vol. 26, pp.143-151.
Underwood J (1998) "Making Groups Work". In *IT for Learning Enhancement,* (Ed.) Monteith M, Swets & Zeitlinger Publisher, pp.29-41.
Vygotsky L (1978) *Mind in society: the development of higher psychological processes,* (Eds. and Trans.) Cole M, John-Steiner V, Scribner S and Souberman E, Harvard University Press: Cambridge, Mass., London.
Vygotsky L (1987) *The collected works of L.S. Vygotsky: Vol.1Problems of general psychology: including the volume Thinking and speech,* (Eds.) Rieber RW, Carton AS, (Translation) Minick N, New York; London:Plenum.
Wardekker WL (1998) "Scientific concepts and reflection", *Mind, Culture, and Activity,* Vol. 5, No. 2, pp.143-153.
Webb N (1985) "Verbal interaction and learning in peer-directed groups", *Theory into Practice,* Vol. 24, pp.32-39.
Webb N (1991) "Task-Related Verbal Interaction And Mathematics Learning In Small Groups", *Journal for Research in Mathematics Education,* Vol. 22, No. 5, pp.366-389.
Webb N (1997) "Assessing Students in Small Collaborative Groups", *Theory into Practice,* Vol. 36, No. 4, pp.205-213.
Webb NM, Farivar SH and Mastergeorge AM (2001) *Productive Helping in Cooperative Groups,* CRESST - National Center for Research on Evaluation, Standards and Student Testing, University of California, <http://www.cse.ucla.edu/CRESST/Reports/TR555.pdf>.
Weeden P, Winter J, Broadfoot P, Hinett K, McNess E, Tidmarsh C, Triggs P and Wilmut J (1999) *The LEARN Project: Learners' Expectations of Assessment for Learning Nationally - Report for the Qualifications and Curriculum Authority,* University of Bristol Graduate School of Education: CLIO Centre for Assessment Studies, <http://www.qca.org.uk/ca/5-14/afl/learners.pdf>.
Wegerif R (1996a) "Collaborative learning and directive software", *Journal of Computer Assisted Learning,* Vol. 12, No. 1, pp.22-32.
Wegerif R (1996b) "Using computers to help coach exploratory talk across the curriculum", *Computers & Education: An International Journal,* Vol. 26, No. 1-3, pp.51-60.
Wegerif R (2001) *Citizenship, ICT and thinking,* <http://137.108.56.58/fels-thinking/FlashPages/MenuPages/MainPublications.htm>.
Wegerif R and Dawes L (1998) "Encouraging Exploratory Talk Around Computers". In *IT for learning enhancement,* (Ed.) Monteith M, Intellect Books, pp.10-28.
Wegerif R and Mercer N (1986) *Using Computer-based Text Analysis to Integrate Qualitative and Quantitative Methods in Research on Collaborative Learning,* CLAC Occasional Papers in Communication, Open University.
Wegerif R, Mercer N and Dawes L (1998) "Integrating Pedagogy and Software Design to Support Discussion in the Primary Curriculum", *Journal of Computer Assisted Learning,* Vol. 14, No. 3, pp.199-211.
Wertsch JV and Stone C (1985) "The concept of internalization in Vygotsky's account of the genesis of higher mental functions". In *Culture, communication, and cognition: Vygotskian Perspectives,* (Ed.) Wertsch JV, Cambridge, MA: Cambridge University Press, pp.162-182.
White B, Shimoda T and Frederiksen J (1999) "Enabling Students to Construct Theories of Collaborative Inquiry and Reflective Learning: Computer Support for Metacognitive Development", *International Journal of Artificial Intelligence in Education,* Vol. 10, pp.151-182.
White BY and Frederiksen JR (1998) "Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students", *Cognition and Instruction,* Vol. 16, No. 1, pp.3-118.

Whitebread D (1997) "Developing children's problem-solving: the educational uses of adventure games". In *Information technology and authentic learning,* (Ed.) McFarlane A, Routledge, pp.13-37.
Wild M (1996) "Investigating verbal interactions when primary children use computers", *Journal of Computer Assisted Learning,* Vol. 12, No. 2, pp.66-77.
Wild M and Braid P (1996) "Children's talk in cooperative groups", *Journal of Computer Assisted Learning,* Vol. 12, No. 4, pp.216-231.
Wills M (2001) *Speech at BETT 2001,* <www.becta.org.uk/news/mwspeech.html>.
Wood D, Bruner J and Ross G (1976) "The role of tutoring in problem solving", *Journal of Child Psychology and Psychiatry,* Vol. 17, pp.89-100.
Wood D, Wood H, Ainsworth S and O'Malley C (1995) "On becoming a tutor: Towards an ontogenic model", *Cognition and Instruction,* Vol. 13, No. 4, pp.565-581.