Supporting Learners in a Remote Computer-Supported Collaborative Learning Environment: The Importance of Task and Communication

David Graves
graves@cs.ubc.ca

Department of Computer Science
The University of British Columbia
Vancouver, British Columbia V6T 1Z4
CANADA

February 25 1998

Abstract

This paper describes novel research in the area of remote Computer-Supported Collaborative Learning. A multimedia activity (*Builder*) was designed to allow a pair of players to build a house together, each working from his or her own computer. Features of the activity include: interactive graphical interface, two- and three-dimensional views, sound feedback, and real-time written and spoken communication. Mathematical concepts, including area, perimeter, volume, and tiling of surfaces, are embedded in the task. A field study with 134 elementary school children was undertaken to assess the learning and collaborative potential of the activity. Specifically, the study addressed how different modes of communication and different task directives affected learning, interpersonal attitudes, and the perceived value and enjoyment of the task. It was found that playing led to academic gains in the target math areas, and that the nature of how the task was specified had a significant impact on the size of the gains. The mode of communication was found to affect attitudes toward the game and toward the player's partner. Gender differences were found in attitude toward the game, perceived collaboration and attitude toward partner.

Important Note: If your copy of this Technical Report does not include a series of sample tests and screenshots at the end of the Appendix, you will need to acquire these via FTP to ftp.cs.ubc.ca. You should log in as anonymous and download the attachment postscript file from the location: /pub/local/techreports/1997/TR-97-19a.ps.gz (note that the screenshots are in colour).

Contents

Chapter 1

Introduction

This thesis presents the results of a study about the effectiveness of a multiuser computer activity for mathematics learning in the intermediate grades. In particular, the study looked at the impact of different modes of communication and different types of tasks on learning and attitude in same-sex pairs. The thesis begins with a discussion of the motivations behind the elements investigated in the study, followed by an overview of the structure of the thesis. As a prelude to the discussion of motivations, a brief introduction to the activity that was used in the study is given below.

1.1 The Game *Builder* – A Preview

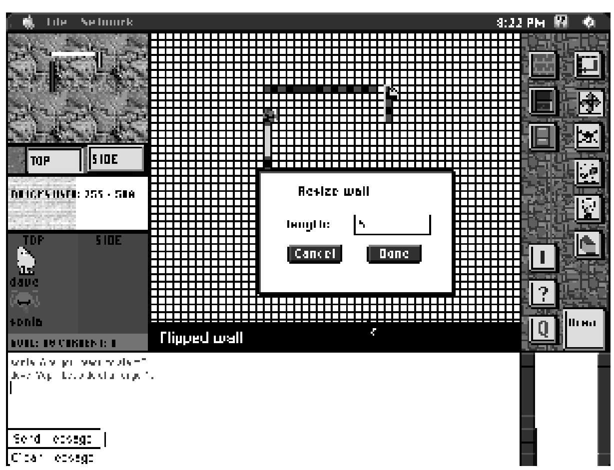

Figure 1.1: Changing the size of a wall in *Builder*

The activity *Builder* was designed and implemented as part of the work contributing to this thesis. *Builder* consists of a series of challenges, each of which asks players to construct a house according to specifications relating to floor area and volume (e.g., "Build a house with a floor area of 80 square units"). It is designed for two players, who work together in a shared graphical virtual space to achieve their goal. Players see the results of each others' actions (such as building a new wall), and can communicate with each other either by spoken or written messages (depending on the mode of the game they are playing). In Figure 1.1 we see that Dave is in the process of changing the size of a wall (Screenshot 3 in Appendix F shows an enlarged, colour version of Figure 1.1). The maximum possible perimeter of the house is limited by the fixed number of bricks allocated for each challenge. Players can add windows and doors to their walls, which frees bricks according to how much area the window or door covers. As with bricks, the number of windows and doors is limited by the set of frame pieces allocated for the challenge. The goal in the design of the activity was to teach mathematical concepts relating to area, perimeter, volume, and tiling of surfaces.

Chapter 4 gives a complete description of *Builder*.

1.2 Motivations

Traditional Distance Education (DE) allows geographically-separated or non-mobile students to take centrally-administered courses. With the increasing use of computers in DE, such as WWW-based projects like WebCT [Gol96], the traditional model has been enhanced in several ways. The learning material, for example, can be easily distributed and updated from the central location, alleviating the inconvenience of postal delays and retrieving material from libraries. Furthermore, in contrast to the regular text materials used in DE, internet-based systems allow material to be of a multimedia nature, including video and sound, and can also include interactive elements providing immediate feedback, such as on-line quizzes [Gol96]. Additionally, these WWW-based facilities allow communication between students and instructors via bulletin boards, email and sometimes real-time chat facilities.

The dual roles of the computer in DE

Reflecting on the qualities brought to DE by computers suggests there are two types of functions involved. On the one hand we see computers used in the organization and presentation of content, while on the other they provide a medium of communication between those involved in the learning. The central questions of this study, which is of an exploratory nature due to the lack of previous research in the field, are prompted by consideration of these two roles. Current computer-based DE courses show significant benefits over the traditional methods, and also employ technology just within the capacity of a typical home computer system. [1] However, as advances in technology allow progressively greater functionality and flexibility in networked computer environments, so the potential role for computers broadens within the DE setting.

Supporting communication

One of the primary issues addressed in this study was the influence of the type of communication on the quality and experience of learning in a remote Computer-Supported Collaborative Learning (CSCL) activity. As stated above, in DE computers may provide

[1] It would be unreasonable, for example, to include a real-time video-conferencing capability in a system intended for use by students with a modem connection.

asynchronous communication in the form of email and bulletin boards [Lev92, Rie92], or synchronous communication in the form of real-time chat [Gol96]. These channels of communication suffer the restriction that they are composed purely of ASCII characters, so a user is unable to draw pictures, for example, to describe what s/he is thinking. Such limited forms of communication may provide supplementary help to students taking a course, but if the goal is for interaction to play a more central role in learning, a richer type of communication may be required. Two fields of research bear upon this issue. The most important of these is Cooperative Learning (CL), since the positive academic and social results in the CL literature [JMJ+81, JJ83, e.g.] provide the incentive to bring collaboration to the DE setting. Secondly, we can consider extensions to DE in light of the parallel field of Computer-Supported Cooperative Work (CSCW), where the goal is to enable a group of remote users to work together on a project in a virtual shared space. Many CSCW projects have developed technology to simulate in-person communication through the use of audio and video channels [IM91, e.g.], based on the assumption that enriched channels of communication will provide stronger support for the cooperative activity. In this study simple written communication was compared with a richer form of communication, which included written and spoken messages, as well as a graphical representation of the other user ("virtual presence" [2]).

Designing content

The second primary issue addressed in the study was the influence of the nature of the task on the learning experience. In computer-supported DE, we have said that the computer is used to present and organize teaching material. Limited interaction and feedback may also be provided in the form of automatically-marked on-line questions that ask rote-style, memory questions about the material. Given that it has been typical for tertiary-level teaching to be concerned more with the content of the material than how it is taught, this type of feedback may be considered appropriate for the university-level students at which projects such as WebCT are aimed. For the current study, aimed at elementary-level students, the role of the computer was extended to allow learners to explore with manipulables in a task in which the learning material was embedded, rather than presented. Again, relevant work in the fields of Cooperative Learning and Computer-Aided Learning (CAL), which is concerned with how to effectively use computers as educational tools, provides motivations and suggestions. For example, researchers in CL, and Computer-Supported Collaborative Learning (CSCL), have placed critical importance on the nature of the task, finding that rote-style tasks are not as effective as less-structured tasks within a collaborative learning setting. In the current study, the role of the task in CSCL was investigated by manipulating the specification of the goal in the activity.

1.3 The Learning Activity

The specific instructional setting used in this study was an interactive, multimedia problem-solving activity (*Builder*), where two learners worked together over a local network from physically-separated computers, communicating with written and spoken messages over

[2] Also referred to as "telepresence" in the CSCW literature.

a network. Mathematical constructs involving area and volume were embedded in the familiar real-world task of building a house. Given the embedded nature of the task and that the activity employed an interactive graphical interface with sound effects, the activity was game-like in nature. This was reinforced by having clearly-specified challenges with scores, and keeping records of high scores between different sessions. The design of such a game involves several Human-Computer Interaction (HCI) considerations, some of which are common to all CAL applications, while some are specific to remote CSCL. In the former case there is the question of what special demands a game or an educational activity places on a computer interface, with the added issue of age-appropriateness. For the latter case, as in the field of CSCW, the demands relate to how an interface should present shared and private work, and how identity, behaviour and communication of remote learners should be represented.

The use of electronic games in learning, either of an individual or group nature, is a relatively new endeavour, and has been the subject of research of the E-GEMS (Electronic Games in Education for Mathematics and Science) group at UBC. E-GEMS' findings have shown potential both for effectively-presented CAL content promoting learning [SK96b], and for successful co-present collaborative play [IBK95]. In these and other electronic game studies, significant gender differences have been found, so in addition to the primary questions concerning task and communication, gender was also considered as an independent variable in the current study.

1.4 Research Questions and Structure of Approach

To summarize the discussion above, the motivating question behind this study was:

> *Can the positive outcomes of Cooperative Learning also be achieved in a remote CSCL environment?*

The research questions addressed in the study were:

- What type of communication is necessary to support remote collaborative learning?
- How does the nature of the task affect the learning experience?
- Are the effects of communication and nature of task the same for male and female learners?

Other research
As discussed above with regard to the motivations of the study, several fields of research bear upon the questions addressed in this study. The fact that so many different fields are relevant is partly due to the multidisciplinary nature of the field of CSCL, and partly because there is so little literature dealing with the specific CSCL model investigated in this study. This lack of research is evident in a recent extensive literature review on educational multiplayer computer games which, despite being 88 pages long, spends only three pages describing the literature on educational multiplayer games [McG96]. Even removing the game restriction, there is relatively little research in the area of remote CSCL, most CSCL studies being in the co-present domain. Therefore, the literature review touches upon many

different but sometimes overlapping fields, drawing from them the most pertinent issues for this study.

The review of other research (Chapter 2), begins with the discussion of Cooperative Learning (CL), CAL and CSCL. It is appropriate to look at these fields together, as CSCL is essentially a marriage of CAL and CL. CL provides a good starting point because it is a well-established field, and many of the motivations, expectations, and assessment techniques, both for CSCL in general and this study in particular, are to be found there. The CAL section introduces the topic of the general use of computers in education. [3] The discussion of CSCL begins by defining four different models within the CSCL field. Much of the work has been done with co-learners in the same room, either at one or more than one computer, which is a substantially different environment to that of the current study. These studies have dominated the field for several reasons: they are easier to do; up until recently the available technology was too limited to support collaboration in rich virtual learning environments; and researchers have implicitly or explicitly believed that the necessary interaction for CL is not supportable remotely.

Since this study uses the setting of a multiplayer game to investigate collaborative learning, the discussion of CSCL is followed by a section specifically dealing with educational games. Although educational computer games are a sub-category of CAL, they are worth considering separately to evaluate their relative efficacy in instructional settings, and because they entail a somewhat different set of design guidelines, which should be addressed in relation to the *Builder* activity. A substantial proportion of the design of educational computer applications (the proportion not directly concerned with the content of the learning material, though the content itself should greatly influence design) is the interface with the user, which falls into the field of Human-Computer Interaction (HCI). The literature review addresses HCI within the discussions of CAL, CSCL and games, which are all fields that involve HCI issues.

The final section considers gender differences, which are a striking feature of many studies in the computer game and CAL literature. The goal in addressing gender differences and other learner characteristics is to ascertain whether the effect of different aspects of the learning environment depends on learner characteristics, which has implications for design.

The current study

Following the literature review, Chapter 3 summarizes the relevance of the previous research to the current study. This includes discussion of the decisions made in the design of the *Builder* activity, and suggestions for hypotheses regarding the manipulations of communication and task structure based on the literature review. The remainder of the thesis presents the current research project, beginning with a description of *Builder* (Chapter 4) and the assessment tools designed for use in the study (Chapter 5), followed by an account of the pilot studies conducted prior to the study itself, which led to modifications to the activity and the planned study procedure (Chapter 6). The study itself and its implications are described in the final three sections: Method (Chapter 7), Results (Chapter 8), and Discussion (Chapter 9).

[3] For the purposes of this paper, the term CAL includes similar labels that have been applied to the use of computers in education, e.g. Computer-Based Instruction (CBI), Computer-Assisted Instruction (CAI).

Chapter 2

Research Context

Much of the promise and excitement of Computer-Supported Collaborative Learning (CSCL) rests with the fact that it combines Cooperative Learning with Computer-Assisted Learning (CAL), two fields which each bring their own different advantages to CSCL. The description of related research will begin with a brief discussion of these two fields and how they come together in CSCL. Following this discussion, other specific research areas of relevance for the current study will be described, specifically computer games and computer-related gender issues.

2.1 Cooperative Learning

Roots of Cooperative Learning

Cooperative Learning is not a new idea, being written about as early as the 1920's; but studies that have evaluated its potential as an educational model have mostly taken place in the last 20 years [Sla80]. Researchers in the field draw upon theoretical perspectives from educational philosophy and social psychology to stimulate and guide experimental work. Allport's work on the positive effect of interaction on race relations, for example, provides motivation for the application of CL in school classrooms, with the goal of alleviating prejudice amongst heterogeneous learners [HH91]. The suggestion that it may also be an effective approach to learning in general emerges out of more philosophical theories such as John Dewey's **situated learning**, in which he saw learning emerging out of purposeful communal activity and transactions with others [Ros92]. This is compatible with the modern view of **constructivism** in which knowledge is seen as emerging out of sociocultural interaction rather than something to be revealed or transmitted [SBM+89]. It should not be implied that these theories reflect the consensus views in the educational psychology literature, as some writers have argued that situated learning and constructivism are academically inefficient and incomplete [ARS97]. Nevertheless, while it is true that collaborative learning models may never entirely replace the more traditional methods of instruction, the positive outcomes in the field suggest that they may play an important role, both academically and socially, in a complete education system.

What is Cooperative Learning?

A central element in the definition of CL is the **reward structure**. The typical model of education is competitive, with evaluation ranking students on a scale of better-to-worse. The continuing dominant presence of competitive-style education in the classroom, despite the grand claims emanating out of the CL literature, is supported by the common-sense belief that competitive-style education is necessary if learners want to be successful in the "real world" [HZD93]. The reward structure in such competitive systems has been labelled **negative reward interdependence**, because one student's success depends on the failure of others [Sla80]. Cooperative Learning researchers such as Slavin and Johnson and Johnson have conducted extensive studies on the differences between competitive, cooperative and individualistic goal structures [JJ74, Sla90, e.g.]. In contrast to competitive goal structures, cooperative goal structures are defined as "a situation in which the only way group members can attain their own personal goals is if the group is successful" [Sla90, p. 13]. Cooperative Learning may or may not involve intergroup competition. The individualistic paradigm is that in which reward does not depend on others at all. The work of Slavin and Johnson and Johnson indicates that **positive reward interdependence** is one of a set of necessary factors for CL. The other factors suggested include: **face-to-face interaction**, **individual accountability**, **social skills training**, and **group evaluation opportunities** [HZD93].

What's so good about Cooperative Learning?
As mentioned above, one of the early goals in CL research was the breaking down of cross-racial prejudice. Several early studies showed positive effects of CL on interpersonal relationships [JJ83, Sha80]. Johnson and Johnson, for example, found that CL led to an increase in positive attitude towards peers and a higher value placed on working with heterogeneous peers [JJ83]. This effect was observed not only in inter-racial groups but also in cross-sex and cross-handicap relationships [JJSR85]. On the basis of such findings, CL has been recommended for classrooms where there is mainstreaming of students with disabilities [Mal86].

In addition to the positive social findings, many studies have also found positive effects on achievement [JMJ+81, ST79, e.g.]. [1] For example, a meta-analysis of 122 studies by [JMJ+81] found that cooperation and cooperation with intergroup competition led to greater academic gain than either interpersonal competition or individualistic goal structures. There have, however, been conflicting findings in the achievement literature, with some findings of competition being more successful [Mic77, e.g.], and debate over the exact conditions in which Cooperative Learning is beneficial. More refined research has suggested that successful CL depends on a number of factors.

What factors are most important in Cooperative Learning?
In several studies Webb has looked at the effect of interaction on achievement, finding that only certain types of interaction, such as giving an elaborate answer to a peer, are positively related to achievement [Web82b, Web82c]. In response to the debate over which type of learners and what type of group composition profit most from CL, Webb's findings suggest that the conflicting data may depend on how often such groups or learners engage in this type of creative, task-related interaction. Homogeneous groups, for example,

[1] Achievement is meant in terms of task-related learning, most commonly measured by pre- and post-tests on the material being covered.

which have been shown in some studies to benefit less from CL, show less of this type of interaction, while heterogeneous groups and extroverted learners show more. In contrast to Webb's results, Cohen and her associates have found that simple frequency of interaction predicted achievement if the problem was inherently a group task and the task was ill-structured [Coh94]. An ill-structured task is one which has no single right answer. Negative support for Cohen's argument is found in [JMJ+81]'s meta-analysis of CL findings, which found that rote-learning tasks had less effect than other types of task on achievement measures. Despite the apparent conflict in the literature, Webb and Cohen's arguments may be compatible in that ill-structured tasks could lead to co-learners exchanging more explanations.

The importance of explanation in achievement results helps illuminate the cognitive processes which make CL effective. In some of Webb's studies, for example, it was found that high-level learners profit more in heterogeneous groups than they do in homogeneous groups. This is hypothesized to be due to there being more of a need for explanation between the high and low-level learners, and that this explanation involves rehearsal and cognitive restructuring in the explainer. Low-level learners also profit more in heterogeneous groups, perhaps because they find it easy to understand the explanations of the higher-level learners. Middle-level learners, on the other hand, profit more in homogeneous groups, which Cohen suggests may be because they don't get a chance to explain when there are higher-level learners in the group [Coh94].

The notion of an "inherent group task" has been studied mainly in terms of giving group members specific roles, or limited tools or knowledge that must be used in combination to solve a problem. Repman, for example, found that groups using roles scored higher on achievement than those without roles, though there was no parallel positive difference shown in sociomotivational results [Rep93].

To summarize the findings of what makes CL successful (in addition to those factors previously defined as the "necessary" components of CL) the literature proposes [2] use of the following ingredients:

- high levels of **interaction**, particularly task-related, explanatory and conceptual interaction;
- **training in cooperative relationships**;
- **small groups** – two or three members work best [JMJ+81];
- enforcing the **group nature** of the task by the assignment of responsibilities, roles or divided resources;
- problems that are **ill-structured**.

2.2 Computer-Aided Learning

In a meta-analysis of 51 studies on computer-based teaching, [KBW83] found that computer-based groups did better overall than conventional learning groups according to academic

[2] Not unanimously, of course.

measures, and that use of the computer substantially reduced learning time. In addition to academic gains, classroom computer activities have been found to lead naturally (i.e., without teacher intervention) to higher levels of student interaction compared to non-computer activities, such as solving a jigsaw puzzle [HSGB82, NC93, JJ86]. Several researchers have observed that students will naturally spend more time, even sometimes outside normal class hours, working on computer tasks than on non-computer tasks [Sol91, Cle81]. This suggests that there may be something intrinsically motivating or engaging about using computers.

There is much disagreement in the literature, however, over the true value of CAL. The research findings have not been universally positive [WM94, e.g.], and the record within actual classrooms has not been as powerful as some of the effects found in the research literature [Rob94]. There are also social concerns that CAL might lead to less interaction with teachers and classmates and hence not allow for sufficient social modelling or building of social skills and healthy social attitudes [JJ86, HDH87].

Proponents of CAL, however, claim that the problem resides in the nature of the CAL applications that have been used and the attitudes towards their use. Hannafin et al [HDH87] argue that the resistance to computers in schools is based on a misunderstanding of the true potential of CAL. On the basis of the studies cited above finding computers led to increased classroom interaction, they propose that the image of the isolated computer user is a myth. Furthermore, they argue that existing commercial CAL applications support the myth that computers are appropriate only for learning low-level skills. Such applications are criticized for emphasizing flashy graphics and sound, while consisting purely of repetitive drilling on rote-style questions [Pap93]. Instead, Hannafin et al argue that CAL has the potential to provide valuable aid in areas that traditional learning cannot provide. The roles suggested include:

- use in one-on-one instruction with a flexible level of difficulty, allowing students to proceed at their own pace;
- use in simulations of scientific phenomena that are impossible to demonstrate in the classroom; and
- use in visualization and manipulation of large sources of data or other information.

Papert and others have also proposed the use of computers as tools for creation, rather than as fixed question-askers. One example of this is is a study by Soloway where students used computers to create their own multimedia packages on a chosen topic [Sol91]. This study led to increased attendance and participation by "problem" students. Another example, described below in the section on CSCL, is the series of Logo studies by Nastasi and Clements in which students were engaged in computer programming exercises.

The use of computers in education includes computer games, which are discussed separately in Section 2.4.

HCI concerns in CAL

The field of HCI is concerned with the usability of software, with the typical goal being to design interfaces that are "transparent" to the user. This traditional approach to HCI is centered around the expert user, aimed at improving productivity in computer-based work.

More recent thinking in HCI has identified the need for different types of interfaces for different users and applications [Car87, e.g.]. For example, when novice users are learning to use an application, the demands on the interface are entirely different than those arising in expert use [TM94]. To support the learning process for novice users, researchers in the field have employed techniques such as scaffolding, in which the full functionality of a system is only gradually revealed, rather than confusing the user with a huge set of unfamiliar functions from the outset [Sol93]. This technique can also be applied to CAL applications, where there is a need to guide learners through the system, at least in the early stages of use. As more is learnt, and the activity becomes more difficult, more tools could become available within the interface. [3]

Some researchers, however, have questioned the relevance of "ease-of-use" and "transparency" within an instructional setting. In CAL, as in electronic games, it is often desirable to have some element of challenge in the activity, which may be incorporated into the interface. Research conducted by Kamran Sedighian using the E-GEMS game *Super Tangrams* has investigated the role of the interface in learning transformational geometry [SK96a, Sed97]. The design of the interface was based on the goal of promoting reflection by intentionally avoiding direct manipulation of objects (as in "drag-and-drop") in favour of manipulation of formal representations of the transformations being used (e.g., the length and direction of a 2D vector). It was found that making learners focus more on the important properties of transformations led to significant academic improvement on pencil-and-paper post-tests [SK96b].

2.3 Computer-Supported Collaborative Learning (CSCL)

The use of computers in a collaborative educational setting is a natural progression from CL and CAL because:

- we want to retain the benefits of CAL, such as its flexible level of difficulty, and usefulness in visualization;
- use of computers appears to promote interaction, compared to other classroom activities;
- CSCL alleviates the concern of computer use encouraging isolation;
- computers provide a means to bring collaborative learning to remote settings, such as in Distance Education.

Types of CSCL

As stated previously, much of the work done in CSCL is within the co-present model, which is easier to study and also provides a natural way to allow communication between co-learners which is appropriate in many learning settings. Co-present CSCL entails a substantially different set of concerns than remote CSCL, since interaction in the in-person

[3] This is potentially confusing, since scaffolding in some applications may consist of removing tools so that the user has to think more. The technique of progressively adding tools might be used when training users on complicated business applications, for example.

model includes the language, facial expressions and gestures of everyday life (i.e., the same type of communication as in CL). In addition to standard co-present CSCL, we identify three other possible types of CSCL. The four types are distinguished by two factors: the type of communication used, and the nature of the shared space. Accordingly, the types are defined as follows.

- **Standard co-present CSCL:** In this model, all learners sit at one computer, sharing one input device. We refer to this as a "literal" shared-space even though the shared-space is electronic rather than physical, because there is literally one screen that learners share. A virtual shared-space is defined as one where multiple electronic environments (physically presented on different screens) are virtually integrated into a single work-space. [4]

- **Multiple-Input CSCL:** This entails use of one computer as in the standard model, but each learner uses her/his own input device. This model is typical in video games. Since learners are co-present, communication is the same as in the standard model. The shared-space is still literal, but may be more complicated than in the standard co-present model because users may be represented separately in the interface. The simplest example of this is the use of multiple cursors differentiated by shape or colour. In video games multiple users are generally represented by avatars. [5]

- **Multiple-Computer CSCL:** Learners work at a separate computers, but are all in the same room. Communication is as in the standard model, but the shared-space is virtual.

- **Remote CSCL:** In the remote model, both communication and shared-space are computer-mediated.

Before describing the studies that have investigated each of these CSCL types, two preliminary issues will be addressed. First we consider the potential advantages and disadvantages in the comparison between CSCL and non-computer CL, which vary according to the type of CSCL being considered. Second, we address HCI issues in designing CSCL applications, which likewise vary according to the CSCL model.

Differences between CSCL and CL

General differences that exist between all types of CSCL and CL stem from the addition of the computer into the learning setting. Since the computer is supplementary to the interactions going on between learners, and can be incorporated into the process as much or as little as desired, there are no clear disadvantages of CSCL compared to CL. The potential advantages of CSCL overlap with the discussion of CAL above, such as the provision of variable levels of difficulty to accommodate learners of different levels. An additional advantage claimed by some researchers on the basis of findings of students staying late or arriving early to work on CSCL projects, is that CSCL instruction places more responsibility for learning on the students [Bat92]. It has also been argued that CSCL ameliorates the

[4] Another way to clarify this is to say that the "virtual" refers to the sharing rather than the space.

[5] "Avatar" is the term used for a graphical representation, often human-like, of a user in an electronic environment.

problem identified in regular CL instruction of students straying onto off-task behaviours, because the computer can be used to set lesson pace [Dal90].

In multiple-computer or multiple-input CSCL, the computer can enhance the collaboration between learners if the virtual shared-space is well designed, because it allows equal access to all participants, the ability to use flexible symbol representation, and easy integration of individual work with shared work. In remote CSCL there is the disadvantage that users can not communicate with each other in the normal way. Given Goldman's [Gol92] findings that physical proximity was related to the highest level of conversation during a CL activity, this absence of physical contact could be detrimental to the learning experience. There may, however, be unexpected advantages to remote CSCL. Batson [Bat92], for example, has proposed that the absence of face-to-face contact, even without true anonymity, encourages free expression in interaction with classmates and teachers. It might also be hypothesized that, at least for younger learners, the novelty of communicating electronically makes the experience more appealing.

HCI issues in CSCL

The majority of co-present CSCL studies involve the pair or group sitting around one computer with one input device, meaning that the interface issues are similar to those in individual CAL. There is no need to consider how the presence of users should be indicated in the virtual space, nor to control simultaneous modification of shared material. In multiple-computer and remote CSCL, special design concerns arise about how to display a workspace that is being modified, perhaps simultaneously, by other users. The remote CSCL designer has to deal with the additional problem of how to incorporate communication between users into the interface. Educational thinkers have suggested that symbols and gestures are social tools for mutual learning [Ros92]. In co-present interaction, these gestures and symbols are drawn from the rich physical and social environment shared by the co-learners. The challenge in remote collaborative activities is to know what symbols to provide and how to best simulate gestures and other forms of communication via electronic media.

Representing the actions of remote users has been a central concern in CSCW research. In the most simple, command-line shared editing systems, the behaviour of other users may be indicated only by the appearance of the text they type. More recent CSCW applications have added icons to represent each user and labelled cursors which can be used in gesturing [GRWB92], with some of the more "high-tech" systems providing full video and audio linkage between participants [IM91, e.g.]. Ishii and other CSCW researchers have emphasized the importance of the shared view. The motivation for the application ClearBoard, for example, which includes super-imposed video footage of the remote user, is allowing workers to follow the eye-movements of their colleagues as they discuss group work [IK92]. The notion of the shared view may also be critical in CSCL applications. Goldman's study of children interacting during a face-to-face collaborative learning exercise, for example, found that when learners were engaged in conceptual conversation they spent more time looking at the shared workspace (in this case a computer screen) than at their own work-sheets, a pattern which was reversed during more routine on-task conversation [Gol92].

Co-present CSCL studies

Based on studies comparing cooperative-style learning using a computer game with competitive and individual learning styles, Johnson and his associates have proposed that CSCL has many of the same relative advantages of Cooperative Learning [JJ86, JJS86]. The cooperative style led to higher achievement on a post-test and was characterized by conceptual and task-oriented interaction. It was also found that those in the cooperative condition had more positive attitudes toward peers; e.g., males were more positive toward females [JJ86]. They argue that the agent of the effect is the positive mood produced by the CSCL activity, which in turn influences motivation and span of attention.

As was the case for CL, it is important to understand the exact conditions under which CSCL is most effective. Are gains uniform across all learner types? Are some tasks particularly suited to the CSCL teaching style? Some studies have addressed these questions, but as yet the variables are not understood as well as in CL. Dalton and associates [DHH89, HH91], for example, studied the relative gains across learners of different academic level. They sought to explain the mixed results in the literature concerning gains made by high-level learners in heterogeneous groups, hypothesizing that the level-of-learner effect may depend upon content of learning material. They found some effects of content, and also that interaction was higher in heterogeneous than homogeneous groups, but they did not find the expected interaction between content and level-of-learner. In Webb's studies of CSCL she found that the key variables identified in her CL studies were not as important within the CSCL setting [Web84].

A study of elementary-grade dyads by Nastasi and Clements [NC93] looked beyond the broad question of the effect of CSCL vs. non-CSCL and considered the effect of varying the situation or task in which the learners were engaged. The main dependent variable considered was motivation, measured according to behaviours displayed by children during the task (e.g., positive self-statements). They compared students working on Logo with students working on more drill-like CAL activities. The Logo activities consisted of students choosing their own problems and defining simple computer programming procedures to achieve them. Logo led to more higher-order thinking, greater motivation and conflict resolution, and more cognitive change. Although it was also found that Logo students experienced more failures than the other CAL groups, this did not lead to more negative self-statements. What appeared particularly important was that the Logo task allowed students to develop divergent ideas, irrespective of whether or not this led to conflict. These results are consistent with Cohen's suggestion that ill-structured tasks are more likely to promote the sort of cognitive activity that leads to effective learning. The data on feedback, however, provided an exception to the other results. The reward system in the Logo activity was largely learner-regulated, which the researchers thought would be an effective way to motivate learners. Instead it was found that the more standard feedback, used in the non-Logo group, where the computer simply announced player success, was more motivating.

Multiple-Input CSCL studies

Bricker and her associates at the University of Washington studied collaboration in a multiple-mouse CSCL activity involving colour-matching and chord-matching. In colour-matching, for example, three learners, each with their own mouse, were asked to manipulate RGB values to approximate a target colour. Some sense of roles was supported in the task in that each learner was responsible for one colour setting, hence success depended on the contributions of all participants. It is worth noting some of the rationale for Bricker et al's choice of the co-present CSCL model, as the points suggest challenges for implementations of remote CSCL. Firstly, they argue that "physically separating students to work individually on computers tends to discourage communication" [BTR$^+$95, p. 2]. [6] Secondly, they question why it is that existing CSCW and remote CSCL applications fail to engage users as much as co-present video games. As Bricker et al point out, the weaknesses in previous attempts at remote CSCL are partly explained by limitations in the existing technology. They also suggest, however, that collaboration is hampered because users are not sharing the same physical view. The challenges for remote CSCL are therefore to support communication sufficiently such that the physical separation does not discourage learners from communicating, and to provide a greater sense that the view of the workspace is shared by the whole group. The results of the Bricker studies were mixed. There was no difference in academic gain between cooperative and individual learning styles, nor on how much students enjoyed the collaborative nature of the task, with some participants saying that they would prefer to do the activity alone. Based on the CL literature it might be hypothesized that the decrease in conflict facilitated by having multiple input devices could lessen the effect on learning. It is possible, however, that the activity was not appropriate for CSCL since there was little need for consultation among learners.

Multiple-mouse collaborative learning has also been investigated by Kori Inkpen of the E-GEMS group using the educational game, *The Incredible Machine* [OL93]. Motivated by initial findings that pairs working together on one computer with one mouse finished more puzzles than either individuals or pairs playing next to each other (but not playing the same game) [IBK95], learners' collaboration was further supported by the addition of an extra mouse. It was found that the use of multiple input devices had a positive effect on achievement, and further that the protocol used to exchange control between learners influenced success in the game [IBGK95].

Multiple-Computer CSCL studies

The design of CSCL settings where learners are working on the same task from different machines requires consideration of several additional design concerns. These include: **view control**, **representation of the user**, **conflict prevention** and the **support of user roles** [SM94]. Implementations of **view control** in CSCL and CSCW applications commonly consist of a shared-space window, and sometimes include areas where the user may edit or view private material. Steiner and Moher's implemention of **view control** in a creative writing CSCL application followed the typical WYSIWIS model, where "What You See Is What I See". **Conflict prevention**, in Steiner and Moher's model, refers to the locking of shared material while one worker is editing them, which is mostly an implementation issue but should also be indicated in the interface. **Support of user roles**, which was not implemented in the creative writing application, is familiar from the discussion of the Cooperative Learning literature, and relates more to the task definition than the interface.

A study on use of the creative writing tool found that partners playing on separate computers in the same room communicated less than partners playing at a single machine. Steiner and Mohler were actually interested in remote collaboration, but put partners in

[6] It is not clear whether this is directed at remote CSCL or the standard single-user CAL model.

the same room to emulate full video and audio communication. They took the results to indicate that even with perfect communication support for learners, the physical separation of the users in a "virtually" shared workspace fails to support collaboration as well as a "literal" shared workspace. Alternative explanations for the results include the fact that the children in the study were very young (kindergarten), had had no previous exposure to multiple-computer activities, and may not have understood the concept of the virtual shared space. Furthermore, in Steiner and Moher's study there was no **representation of the remote user** to signify the other user in the virtual environment. It could also be argued that mediating the communication through the computer might allow learners to interact more easily while remaining focussed on the activity.

Remote CSCL studies
Scardamalia et al. designed a computer-supported collaborative tool named CSILE (Computer Supported Intentional Learning Environments) which allowed groups of learners to build and use a database of domain-related learning material – including assignments, reference materials, and notes between users expressing learning goals and open questions [SBM+89]. The intention was to create a knowledge-building community, resting on the constructivist belief that knowledge is constructed rather than revealed, and that this construction emerges out of sociocultural interaction. Projects such as CSILE can be distinguished from remote CSCL projects like *Builder* in that CSILE is asynchronous rather than synchronous. In CSILE the computer performs the roles of multimedia bulletin-board and library, both of which are tasks that computers excel in. Given the fact that network speed is not an issue in asynchronous settings, the techniques of CSILE can easily be applied to current Distance Education endeavours [Uni97, e.g.].

Distance Education (DE)
DE systems, as discussed in the Introduction, are also examples of remote CSCL. Developers of these systems argue that networked computers provide a link between users leading to a community of learners, rather than the isolation in traditional distance learning systems [Gol96]. The real-time interaction available through chat facilities in such systems differs from more advanced collaborative systems such as those seen in CSCW [IK92, e.g.] because they are not integrated into a more general shared environment. It is not possible, for example, to "point" at something, or to drag a diagram from a virtual text-book into the shared-space. Furthermore, the content of a DE system generally does not involve group tasks, indicating that cooperation between learners is not a central concern. Riel's case study on *Learning Circles* [Rie92], however, was intended as a collaborative learning project, using the same asynchronous tools found in DE and CSILE. Riel's position is the antithesis of the more common approach that strives to simulate in-person collaboration in the remote setting. She argues that asynchronous email-based systems allow each learner temporal flexibility in presenting divergent ideas, in contrast to the time-pressure in synchronous collaboration, which may result in some learners being left out.

2.4 Computer Games

Computer games have had a mixed reception in terms of both their social impact in general and their perceived potential in education. The active violence (as opposed to the passive violence in television) and gender stereotyping found in many popular games has been criticized as psychologically harmful [Pro92], with some studies showing relationships between violent video games and violent behaviour [SW87, e.g.]. Provenzo has also questioned whether electronic games are useful in educational settings given their "programmed" nature – i.e., they come constrained by the creator's world view [Pro92].

Many researchers in the fields of HCI and CAL, however, have looked at the impact of electronic games with great interest. Malone, one of the earliest computer game researchers, has suggested that the intrinsic motivation [7] stimulated by games may be replicable in non-game applications [Mal82]. Even elements of computer games that appear contrary to what is desirable in a non-game application (such as challenge and difficulty to master) may be relevant because these factors are based on increasing levels of difficulty, which can be useful for user-tailored applications in any domain, and can contribute to sustaining user motivation and engagement. Malone's investigations have attempted to establish what it is that makes computer games so motivating. He found that the presence of a goal was the most important factor in determining the popularity of a game. This was followed by automatic scorekeeping, audio effects, randomness, and speed. More formally, Malone's findings can be summarized into the following list of requirements for a motivating game.

- **Clear goal:** This should include performance feedback about how close the player is to achieving the goal.
- **Uncertain outcome:** This is often achieved via the use of variable difficulty levels and successive layers of complexity.
- **Fantasy:** Does the game employ emotionally-appealing fantasies? Does it embody metaphors from physical or other systems that the user already understands?
- **Curiosity:** This is achieved by providing sufficient informational complexity, including audio and visual effects, and elements of randomness.
- **Progressively-revealed information:** Does the game introduce new information as players advance to higher levels?

To this list other researchers have added: novelty, complexity, surprisingness, illusion of control, goal formation, and competition [TM94].

Computer games in education
Given the motivational aspect of games, it is tempting to explore their potential in education, especially for younger students. The findings of academic success of games have been mixed, but show enough promise to warrant further investigation. In a review of 68 studies of games in the classroom, for example, Randel reports that most studies found game-based learning equivalent to or better than traditional learning according to academic

[7] An intrinsically motivating activity is one which does not depend on reward from outside of itself.

measures [RMWW92]. Randel also suggests that games may be more useful in mathematics and science instruction than in the social sciences or arts, and that games may help with retention of rote-like material. As might be expected, students were more positive about games than traditional learning activities [RMWW92]. There are two comments to make about this comparison of games with traditional learning. Firstly, the quality of the game must be questioned. Many computer games are made by commercial companies where the main goal is to sell the product, and critics have argued that the limited learning activities that they offer are generally "tacked on" to the game, rather than being an integrated part of it [Bro93]. Secondly, as [KP95] have suggested, it may not be appropriate to compare electronic games with traditional learning. The easiest educational goal to achieve with electronic games is stimulating interest in the topic. In well-designed games, students may test hypotheses, develop problem-solving strategies and increase their understanding of complex concepts [SK96b, e.g.], but the most effective use of computer games occurs when their use is supported within a broader instructional environment.

Multi-User Dungeons (MUDs)
MUDs, which have appeared in countless forms on many internet servers since their inception in 1980, represent one of the earliest attempts at multiplayer games. MUDs are based on fantasy and role-play, and were not originally designed to be educational, but recently some researchers have designed MUD-like learning environments, seeing great potential for exploratory and creative educational activities [Res92]. A large part of the attraction of MUDs is the collaborative nature of the play, though until recently the domain has belonged chiefly to computer-oriented males [KR93]. The limited popularity of MUDs may be due to the fact that most versions of the game are played entirely on the command-line. Interestingly, attempts to move MUDs into the interactive graphical domain have met with limited success [MF91, e.g.], but this may be due to current limits in technology or the imagination of the creators, and interest appears to be steadily growing.

The broader game *Island* within which the *Builder* activity is set is an example of a graphical, educational MUD. *Builder* itself, which in the study was played without reference to the broader context of *Island*, is not typical of a MUD environment because the intention is for players to undertake specific tasks which begin and end within one session of play.

2.5 Gender Issues

The gender imbalance identified above in relation to MUDs has also been observed in other types of computer games and in the recreational use of computers in general. Surveys quoted in the popular media, for example, have estimated that 75% of video games are bought for boys [Per94]. Other studies referenced in the press have found that women tend to use computers for specific tasks and are more likely to ask for help using computer applications, while men are more likely to "muck around" with computers [Bul94]. This attitudinal difference is argued to impact on perception of mastery over computers and thereby proficiency in use [Bul94].

Research on children's attitudes to electronic games has also found gender differences. In general, boys appear to be more interested in electronic games than girls, though this may be attributed to a male bias in the design of games, many of which feature male-oriented themes of action and violence, and project stereotyped images of strong males and helpless females [Pro92]. The imbalance is self-perpetuating, since the attraction of young males to computers via these games helps provide the next generation of the software developers, thereby maintaining the male dominance of the industry. Studies have shown that the gender differences in attitude towards computer games depend on the type of activity. Boys tend to enjoy competitive games where the aim is to get the highest score [OKdV96], whereas girls prefer games that revolve around relationships [SU97], and are more likely to want to play games that involve collaborating with others [IUK+94b].

In the E-GEMS study on multiple-input CSCL cited above, gender differences were also found in collaborative game-play [IBK95]. It was found that collaborative play had a greater effect on achievement in girls than boys, and that the effect of different interface styles depended on gender. Specifically, a "give" style, where one player passes control of the cursor to the other player, was more effective for girls playing together, while a "take" style, in which the user not currently in control of the cursor can take control, was more effective for boys [IBK95]. Relatedly, in a study looking at control over speed of presentation of lesson material, Dalton [Dal90] found that males preferred it when they were in control, while females preferred it when the speed was set.

Chapter 3

Situating the current research

3.1 Research Focus

To reiterate the goals of the current study, the following broad issues were considered:

- Can the positive outcomes of Cooperative Learning also be achieved in a remote CSCL environment?
- How can we best facilitate collaboration within the remote CSCL setting?

The prior question was addressed partly by a comparison of academic tests between learners who used the CSCL activity, *Builder*, and a no-instruction control group, which allowed us to ascertain whether the activity led to any learning improvement. More generally it was anticipated that the experience of designing the activity and observing students working with it would provide helpful insights on the potential for remote collaboration. In terms of the experimental design and what we can statistically conclude from the study's results, however, the study was focussed more on the latter question than the prior. There was, for example, no attempt to compare the cooperative structure with a competitive or individualistic structure, nor was there any comparison between remote and co-present CSCL styles. The literature reviewed in the previous chapter provides several justifications for the approach taken. To begin with, based on the large number of studies comparing cooperative, competitive and individualistic learning styles [JMJ+81], we can consider cooperative learning established as an effective method of instruction, and turn our focus more to exactly how it is used. The literature also suggests that direct comparisons between learning styles, such as CSCL versus traditional classroom learning, may be ill-advised because of the multitude of additional factors in a CSCL environment that are difficult to control for, such as the effect of interactive graphics on enjoyment [Bat92]. This view was echoed in [KP95]'s suggestion that it is inappropriate to compare games to other instruction because their intended use should be as a stimulating supplement to other learning, perhaps reinforcing material already learned or prompting interest in further investigation of a domain. Furthermore, recalling the introductory discussion regarding Distance Education, it is inappropriate to compare remote and co-present CSCL if the goal is to bring collaboration to a DE setting. Rather, assuming that learners do not have the option of in-person collaboration, how can we best support their collaboration remotely? This approach has practical applications because it generates potential guidelines for the development of DE systems.

The literature review also provides justification for a study explicitly concerned with remote collaborative learning in that very few previous studies have been performed. This suggests that the granularity of investigations in the area should be larger than that, for example, in a current study in the field of CL. We should also expect some difficulty in the interpretation of results due to the involvement of factors whose effects are as yet poorly understood (e.g., the effect of interactive graphics on motivation). The study should therefore be seen as exploratory in nature, and used as a starting point for later, more refined inquiries.

3.2 The Learning Setting

The software developed for this study was a multiplayer, multimedia game, which allowed players to communicate using real-time spoken or written messages while performing the task of building a house. The game context was considered appropriate as a learning style for the target age group because of the intrinsic motivation games can provide [Mal82], and because of the success found in game studies by the E-GEMS group [SK96b, IBGK95]. Previous E-GEMS studies investigated individual games and single-computer multiplayer games where the students were co-present and working at a single computer [Sed97, IBGK95]. The current study expands on this work by turning to an investigation of multiple-computer, multiplayer games where players are physically separated.

Previous research, particularly in the fields of CL and educational games, suggests many factors that should be considered in the design of the activity. Firstly, each of the key ingredients identified in the CL literature will be addressed in relation to *Builder*.

- **Positive reward interdependence:** As discussed previously, positive reward structures are those in which an individual's success depends on the success of the group. In the *Builder* activity, two players share building materials and work simultaneously on the same house. Hence all actions of an individual affect the game status of her partner (i.e., the amount of remaining materials and how close the pair are to the goal). Within the game, scores are assigned to pairs, so there is no concept of individual success. [1]

- **Face-to-face interaction:** All remote CSCL and CSCW depends on the assumption that face-to-face interaction is not essential in a literal sense, but that it can be approximated by the use of appropriate computer technology. This issue was addressed in the *Builder* study by manipulating the mode of communication available between the two computers. As described further below, one mode allowed only written communication, while the other provided a more enhanced form of communication, which included speech.

- **Individual accountability:** This was included in the *Builder* study through the inclusion of academic pre- and post-tests which students completed on an individual

[1] The fact that *Builder* keeps a record of high scores that are displayed at the beginning of each session suggests that, strictly speaking, the model used was that of cooperation with intergroup competition. There was, however, little emphasis placed on trying to beat the high scores.

basis. Players were reminded before they started the activity that they would be sitting a post-test, and that their scores on the test would be individual.

- **Small groups:** *Builder* is played in groups of two, a size that has been found to be effective in CL.

- **Roles and/or divided resources:** This is one element of CL that was not incorporated into the *Builder* activity. We felt that it would be worthwhile to see whether or not, and how, learners collaborated in the absence of such techniques, which aim to structure and enforce the collaboration. The danger of omitting roles is that one learner will do all the work, and there is nothing to prevent this from happening in *Builder*. This problem might not, however, be as prevalent within the game setting, since it is more likely both players will *want* to contribute for the sake of their own enjoyment. It might also be hypothesized that a multiplayer, multiple-computer format provides greater potential for each user to contribute as much as they wish, rather than a situation in which only one learner can be in control of the collaborative piece of work at a time (as in single-input CSCL). A final point on this issue is that the virtual shared workspace of *Builder* makes it impossible for learners to keep their work separate, given that they work at all.

- **Social skills training and group evaluation opportunities:** Training was incorporated to a limited degree in the study by instructing players how to communicate with each other and encouraging them to do so as much as possible. There was no scope in the study for more formal social skills training, nor was there an attempt to employ formal evaluation techniques. The need for the latter is partly alleviated by elements in the game setting, such as the presence of goals and previous high scores, according to which players can judge their performance.

- **Ill-structured tasks:** This was addressed in the study by manipulating the type of task that pairs were assigned within the *Builder* activity. Some pairs worked on a task with a very clearly defined goal, while for others the goal was stated more generally. This is described further under "Variables Investigated" (Section 3.3), and in the succeeding chapter describing *Builder*.

Secondly we will consider each of Malone's criteria for a motivating game with respect to *Builder*.

- **Presence of a goal:** Area and volume goals, with scores providing constant feedback on performance during play, are central to *Builder*.

- **Uncertain outcome:** Malone suggests that uncertainty can be achieved through the use of variable levels of difficulty. This was incorporated in *Builder* by providing five increasingly difficult challenges for players to attempt.

- **Fantasy:** *Builder*'s use of the real-world metaphor of building a house is a good example of what Malone means by "fantasy".

- **Curiosity:** This is related to the informational complexity of the game and is incorporated in *Builder* through the use of interactive graphics and sound. The novelty of allowing players to view a 3-D representation of the house they have built also contributes to satisfying this criterion.
- **Progressively-revealed information:** This was not appropriate for the version of *Builder* used in the study since there was only a single 30-minute session of play.

Finally, to revisit the criticisms in the literature of flashy, low-content educational games, the design of *Builder* balances entertainment and education by fully embedding the learning elements within the central activity of the game.

3.3 Variables Investigated

To better understand the role of some of the key factors in remote CSCL, two variables within the *Builder* activity were manipulated in the present study. The first of these was the interaction between the users, which was addressed in terms of both the medium of communication between the users, and also how they were represented in the virtual space. Communication between co-learners is thought to be the key to many of the gains seen in Cooperative Learning: "Increased verbalization forces cognitive restructuring and re-processing of information, as well as rehearsal and practice of relevant information and skills" [HLKM92, p.258]. On one hand, based on CL researchers' emphasis on face-to-face interaction, it might be argued that in a remote, computer-supported setting, communication should come as close to face-to-face as possible. Indeed, this is the assumption made in video-supported CSCW applications which project images of remote users as they communicate with each other over the network. On the other hand, some studies reviewed in the previous chapter [Rie92, e.g.] have argued that removing face-to-face communication can have positive effects (e.g., by allowing temporal flexibility in the learners' exchange of ideas). It is evident that much research remains to be done on what type of communication is necessary or ideal in the context of educational environments for children. Most current DE systems employ only written communication between users, either synchronous or asynchronous. This study added spoken communication to this model, a closer approximation to an in-person setting, to ascertain whether the learning and attitude outcomes were different from those in the written communication mode. Apart from communication *per se*, face-to-face interaction entails an awareness of the presence and behaviour of other learners, which becomes a non-trivial issue when learners are interacting remotely. Again this is an issue that is addressed in the CSCW literature, though one whose solution is particularly domain-dependent. As a first step in considering how users should be represented in a educational activity for children this study investigated the effect of adding graphical representations of users. To sum up, the present study compared two communication modes: written communication and "enhanced" communication, within which players could both speak and write, and also saw graphical representations of each other.

The second variable manipulated was the nature of the task. Cooperative Learning studies have looked at the influence of content [DHH89], nature of instructions [HLKM92] and structure of task [Coh94] on learning and other outcomes. In our study, the role of the task was investigated by varying the goal specification: some subjects were given a very specific goal (a magnitude of area or volume), such that they would know immediately when they had reached it, while others were given a general goal which was partly open to their interpretation (they were told to maximize the area or volume). [2] The literature review provides conflicting suggestions for hypotheses regarding the comparison between these two conditions. On one hand, Cohen's [Coh94] work, which provided the initial motivation for the different goal modes examined in this study, suggests that ill-structured tasks lead to more successful learning. We might predict that learners have to communicate more when trying to maximize the area or volume, and that this may lead to more reflection or cognitive change. On the other hand, those given a specific goal know immediately when they have performed well. Based on the game literature [Mal82, e.g.], and also the findings of Nastasi and Clements [NC93] that direct external positive feedback from the computer was more motivating than trying to promote an internal sense of motivation, we might predict that the more direct feedback of the specific goal outweighs the value of the more ill-structured task.

Finally, much research on Cooperative Learning and other forms of learning has addressed how the characteristics of the learner affect what is achieved through the learning activity [Web82a, JJSR85, e.g.]. One much researched characteristic is gender, which is particularly relevant in the fields of CAL and CSCL due to the apparent gender differences in interest in, and mastery of, computer applications [IUK+94a, Per94, Bul94]. Thus, in the present study, we had learners play *Builder* in same-sex pairs, allowing us to compare outcomes for boys and girls. The motivation for this comparison is not to judge achievement according to gender, but rather to be sensitive to the fact that some environments may favour particular types of learner, which has to be addressed in instructional design.

3.4 Outcomes Measured

The primary outcomes measured in the study were academic improvement and sociomotivational attitudes, specifically, perceived collaboration, persistence of interest in the game (a typical measure of motivation) and attitude toward the task and the partner. These dependent variables are borrowed directly from the field of CL, where most studies have looked at either one or both of these areas. Pre- and post-tests on the mathematical concepts embedded in the activity (e.g., area, volume, perimeter, tiling of surfaces) were used to assess academic improvement. Following the example of previous E-GEMS studies [IBK95, e.g.], performance in the game was used as an additional measure of achievement. The sociomotivational outcomes were evaluated with an attitudinal questionnaire. While there are alternative methods to assess sociomotivational outcomes, such as the behavioural observation used by Nastasi and Clements [NC93], the questionnaire is perhaps the most common method and is slightly less open to the problem of researcher bias.

We were interested not only in the effects of communication, goal and gender on achievement and attitude (i.e., the "main effects"), but also in the effect of different combinations of these variables (i.e., the "interaction effects"). For example, it may be that the value of a communication or task mode is gender-dependent. Looking at the interrelation of task and

[2] Further details are given in Section 4.2.

communication is also interesting. The fact that there are advantages to either of the task styles, for example, prompts the question of whether the most appropriate goal specification might depend on the employed mode of communication.

Chapter 4

Builder

4.1 Overview

Work on the CSCL activity *Builder* has been done within the context of the E-GEMS (Electronic Games for Education in Math and Science) group at the University of British Columbia. E-GEMS is a collaborative effort involving computer scientists, mathematicians, educators, professional game developers, classroom teachers and students, aimed at motivating children to learn and explore mathematical and scientific concepts with the aid of computer games. Among E-GEMS' current projects is the multiplayer game *Island*, in which the *Builder* activity is set. *Island* is a graphical, educational Multi-User Dungeon (MUD) in which players solve mathematical puzzles to collect materials to build houses on the island.

Island was created for the Macintosh platform in the programming language C++ using a client-server model. It runs over an AppleTalk network, using the NetSprocket library and OpenTransport. The 3-D renderer within the *Builder* activity uses Macintosh's QuickDraw3D library. Further details on the tools used and programming credits are given in Appendix A.

Builder allows two players to design a house using various 2-D layouts and view it in 3-D. In the 2-D design phase players can switch between top-view and side-view, placing and resizing walls, windows and doors. When players have finished designing their house, or at any intermediate point, they can enter the 3-D view and navigate around to inspect the rendered representation of their house.

The following discussion of the views in *Builder* should be read along with inspection of Screenshots 3, 4 and 5 in Appendix F. The screenshots show Dave's screen as he plays *Builder* with Sonia, who is in another room playing on a different computer. In Screenshot 3 (Top-View) they are laying out walls to define the rooms of their house. In Screenshot 4 (Side-View), Dave is placing a window in one of the walls. The final shot is Dave's view in the 3-D rendered environment. He can see Sonia's avatar moving in front of the walls of their semi-constructed house. In the version of the game they are playing, Sonia and Dave can communicate via either spoken or written messages. The two white boxes at the bottom of the screens are for sending and receiving written messages. Microphones and speakers are used for the input and output of spoken messages.

4.2 The Task

We will begin the more detailed discussion of *Builder* with a description of the content of the activity, which centers around the task of building a house. A single session of playing the game may consist of several separate "challenges", with each challenge corresponding to the building of one house. In the target age group (grades 5-9) the mathematics syllabus includes material on calculating perimeter, area and volume for various shapes, as well as tiling of surfaces. These concepts are embedded in the *Builder* activity, as described in detail below. The learning target is that conceptual understanding will be improved in the following areas: addition and subtraction of areas and volumes; tiling of surfaces; and the relationship between perimeter and area (e.g., the fact that a square encloses a greater area than a thin rectangle of the same perimeter).

As one of the goals of the current study was to investigate the role of task in a CSCL setting, two task modes are implemented in *Builder* which differ in terms of how the area or volume target is stated. Players in *specific goal mode* (SGM) are given a numeric target which they have to reach exactly for successful completion of a challenge, while players in *maximize goal mode* (MGM) are instructed to build a house with the largest possible area or volume given the materials available. In the latter condition it is entirely up to the players to decide when they have satisfied their goal. [1] In the former there is direct feedback indicating when the challenge has been sucessfully completed. The two modes are otherwise identical. Both consist of five challenges, with the first three measuring house size in area and the other two in volume. The size of the house that can be built in each challenge is constrained by the available resources. For walls the limit is set by the number of bricks allocated to the players at the start of each challenge. For windows and doors the constraint is not in regard to the material used for the window or door itself, but on the wood pieces that are used to frame the window or door. As for the bricks, a limited supply of these pieces (which are of various length) is allocated at the start of each challenge and shared between the two players. Inserting windows and doors in the walls frees underlying bricks according to the surface area covered by the window or door, which in turn is set by the horizontal and vertical frame pieces they choose. Players are further constrained by a maximum allowable floor area per room so that they cannot simply build one big room.

The mathematical concepts are integrated into the activity in the following ways.

- *Relationships between perimeter, area and shape:* When arranging walls to make a room in Top-View (as described below in Section 4.4.3) players can create houses of greater area [volume] if the room is approximately a square [cube]. For example, two walls of length 20 with two walls of length 4 can define an area of 40 square units (or 72, depending on placement), while four walls of length 10 can define an area of 80 (see layouts in Figure 4.1). Because of the smaller perimeter of the latter room it uses fewer bricks (80 less if the walls are of height 10, which is the default). When the house consists of more than one room the optimization becomes more complicated. Similarly, in Side-View, players should discover that choosing vertical and horizontal frames of roughly equal size will release many more bricks than choosing

[1] Players in MGM are given an indication of expected performance by the high-score records (described in Section 4.4.1)

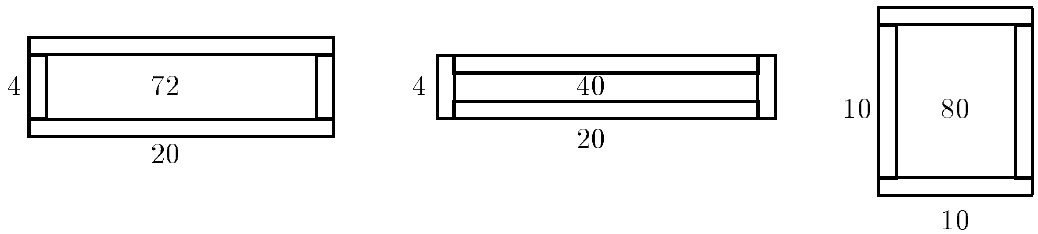

Figure 4.1: Three examples (a,b,c) of room layouts

disproportionate frames. For example, two 1x9 windows release 18 bricks while one 9x9 plus one 1x1 release 82 bricks.

- *Addition and subtraction of areas and volumes:* When placing walls in Top-View to define a house, the width of the walls is not included in the calculation of floor-area and volume; i.e., the measurement is of internal area. For example, if players make a house by laying out 4 walls of length 10 as indicated in Figure 4.1 *c*, the floor area will not be 100 (as might be expected from the calculation 10x10) but 80 square units. It is almost impossible to ignore this issue when aiming for a specific area. Though it is less obvious, this concept is also involved in the placement of windows and doors in Side-View. For example, if the player makes a square window with sides of length 5 in a square wall with sides of length 10, the surface area of the wall *not* covered by the window can be calculated by subtracting the area of the window from that of the wall (i.e., [10 x 10] – [5 x 5] = 75 square units). Calculating the remaining area like this is useful because it allows the user to know how many bricks the wall requires. Since the bricks are square, with sides of length 1, the number of bricks used is in fact the same as the area of the wall not covered by doors or windows. As these calculations are not strictly necessary in attaining the goals, they are performed by the program and displayed to the user beneath the currently active wall in Side-View (described further under *Side-View* in Section 4.4.4).

- *Tiling:* The concept of tiling is naturally embedded in the use of bricks in the activity. When a player needs a certain number of bricks to make a new wall or enlarge an existing wall, s/he is motivated to calculate the door or window dimensions needed to free this number of bricks. To make tiling of surfaces a stronger focus of the activity, harder challenges could be introduced with bricks of varying dimensions. Given the short playing time for the current study we chose to keep bricks as 1 square unit to keep the tiling calculations simple.

Based on previous research, it is not clear which of the two task modes should be expected to generate greater understanding of these concepts. In SGM, we might expect that at least limited awareness of the concepts is assured assuming that the player approximates and then achieves the numeric target. However, once the target is reached players are unlikely to experiment further with alternative strategies. On the other hand, in MGM the degree of exposure to the concepts is set more by the ambition and curiousity of the players.

Players might settle for a relatively low area or volume and hence discover less than those in SGM. Alternatively, they might work exhaustively at the job of maximizing the size of their house, and hence discover more than is needed to meet the specific targets in SGM. Thus the relative utility of SGM vs. MGM in enhancing learner awareness of particular mathematical concepts remains an empirical question, and no specific hypotheses are provided regarding the superiority of ine over the other in the present context.

4.3 Communication

Two forms of communication have been implemented in *Builder*: written and spoken. Written messages require typing in a Send-Message box at the bottom of the game window, and then clicking on the "Send Message" button. All messages, including one's own, appear in a scrollable Receive-Message box also at the bottom of the game window. The message boxes can be seen in all of the screenshots in Appendix F. Spoken messages require the user to hold down the `Control` key while speaking into the microphone. Sound compression allows speech to be transmitted with minimal delay.

For research purposes, two modes of communication within *Builder* are defined. The *basic communication mode* (BCM) allows only written communciation. The *enhanced communication mode* (ECM) allows both written and spoken communication as well as an element of "virtual presence". The term, virtual presence, refers to the simulated presence of other participants in shared virtual spaces, and has been a focus of several CSCW applications [IM91, GRWB92, e.g.]. In *Builder*, virtual presence is implemented in two ways. First, within the 2-D building environment a small icon representing each player appears on top of the wall s/he is working on (see Screenshot 3). Second, when both players are exploring the 3-D model of the house they have constructed, each of them can see the other's avatar moving around (see Screenshot 5). Due to constraints on time and subjects, it was not possible to examine the effects of speech and virtual presence separately. The relative lack of previous research, however, makes a coarse-grain comparison such as this an excellent starting point upon which to base more refined future investigations.

4.4 Sequence of Play and Interface

This section describes the sequence of screens and components of the interface for each screen and should be read along with inspection of the screenshots (Appendix F). Dave is the player whose screen we are looking at; Sonia is the other player. The message windows are present throughout all *Builder* screens.

4.4.1 *"Challenge-Selection" (Screenshot 1)*

This screen shows five buttons marked "Challenge 1", "Challenge 2" etc. The main purpose of the screen is to allow players to choose which challenge they want to do, and to provide feedback on challenges already completed. If Dave tries to select a challenge before Sonia has entered the game, the message "waiting for partner" appears. [2] Once both players have

[2] It is possible to turn this off either at compile-time or run-time to play the game in one-player mode.

entered the game, either player can choose a challenge, which results in both players being sent to the Challenge-Info screen described below. If both players choose a challenge, the choice of challenge is determined by which message reaches the server first. Two lists of scores are provided on the Challenge-Selection screen. The first column (adjacent to the buttons) gives the scores for the current session, while the second gives the record scores for each of the challenges. [3] There are two condition-specific differences for this screen. First, the scores are different according to the goal condition. In MGM (*maximize goal mode*), the scores indicate the area that was obtained for each challenge – hence higher scores are better than lower scores. In SGM the scores refer to how many bricks were used to successfully complete each challenge – hence lower scores are better. For SGM there is definite success or failure, and the brick-count is only recorded if the challenge has successfully been completed. For MGM the area is always recorded if it is non-zero; i.e., there is at least one enclosed room. The second condition-specific difference is that players in ECM (*enhanced communication mode*) can choose an icon to represent themselves from this screen. An additional button (labelled "Choose icon") is provided for this purpose. Players who do not select an icon before choosing a challenge are assigned default icons (chess pieces). Clicking on the "Choose icon" button displays a simple auxilliary screen showing the icons to choose from. The icons are 32x32 pixel images which were downloaded from the WWW. Clicking on one of these images sets the player's icon and returns the player to the Challenge-Selection screen, with the "Choose icon" button no longer visible. The only variation is if the icon has been picked already, in which case a message asks the player to choose again.

4.4.2 *"Challenge-Info" (Screenshot 2)*

This screen presents the goal and available materials for the selected challenge. The center of the screen displays the challenge goal and states the maximum floor area per room constraint. For example, for Challenge 1 in SGM the screen displays the following:

```
Challenge 1 Info.  Your goal is to build a house with area:  80 square
units.  Maximum floor area per room:  100.
```

Whereas for Challenge 1 in MGM the screen displays:

```
Challenge 1 Info.  Your goal is to make the largest possible house.
(Size is in area.)  Maximum floor area per room:  100.
```

At the bottom left the current number of bricks in the brick store is indicated. The sizes of the available frame pieces are listed at the bottom right. Most of this information is also available in Top-View, with the exception of the list of frame pieces which is shown only in the dialog box when players set the frames.

This screen serves the additional purpose of providing a signal that the challenge is about to begin, which is especially valuable for the player who did not choose the challenge. Clicking anywhere on the Challenge-Info screen will send the player into Top-View, which is the main building screen. Players can spend as long as they wish at the Challenge-Info

[3] As the record scores are stored in a file, the records date back to whenever the file on the machine was last moved or deleted.

screen, irrespective of the partner's actions (e.g., Sonia can start building while Dave is still looking at this screen). It is also possible to view the Challenge-Info screen at any time by clicking on the question-mark button from the Top-View screen.

4.4.3 *"Top-View" (Screenshot 3)*

Manipulation of walls
The central area, marked with grids, is for laying out the walls of the house. Both players can be active in the region simultaneously, but a particular wall can only be manipulated by one player at a time. Dave's currently active wall is indicated by a yellow highlight. Sonia's currently active wall is coloured green. In ECM, players' icons are displayed on their active wall as a further indication of ownership. The following actions can be performed on walls in Top-View:

- Walls can be **created** by clicking the wall button on the right-hand button-bar (this button is labelled in Screenshot 3). This causes a wall 5 bricks long by 10 bricks high to appear in the lower left of the central area. Until the player that created the wall selects another Top-View object, only s/he can manipulate the new wall. Under the wall button on the button-bar are the window and door buttons, which are greyed-out and disabled in Top-View, producing an alert sound if clicked.

- Walls can be **selected** by clicking on them. If the clicked wall is currently selected by the partner, the game displays the message "Partner's wall" and plays an alert sound. Successful selection is indicated by a yellow highlight appearing around the wall, and the previously selected wall is de-selected.

- A selected wall can be **moved** by clicking and dragging with the mouse. Upon release, the walls snap-to-grid, allowing ease of alignment with other walls for the user and simple, whole-number area calculations for the program.

- A selected wall can be **flipped** 90 degrees by clicking on the flip button on the button-bar. The flip, resize and delete buttons appear in that order on the far right column of Screenshot 3.

- A selected wall can be **resized** by clicking on the resize button on the button-bar. This produces a dialog box displaying the current dimensions of the wall, which can then be reset by the player to the desired size. The height of a wall can only be set in the last two challenges, which are concerned with volume. For the first three challenges, walls are fixed at a height of 10 bricks, and the height is not displayed in the resize dialog box. Resize requests that call for more bricks than are currently available in the brick store are disallowed and the player is notified by a message displayed to the screen. Resize requests that would leave windows or doors not completely contained by the wall are also disallowed. Otherwise the wall is redrawn with the desired dimensions. Height changes are not visible from Top-View, but can be seen when the player enters Side-View.

- A selected wall can be **deleted** by clicking on the delete button on the button-bar. The wall's bricks and any window or door frames used on the wall are returned to the store.

Feedback and statistics

- **Shadow box:** The top left box of the Top-View screen contains a shadow image of the top view of the working area, with the walls displayed in miniature. Mirroring the colours in the main building area, Dave's active wall is yellow and Sonia's green. This box is also visible in Side-View, where it keeps the player aware of the current layout of the house. In both views it provides a means of selecting walls.

- **Statistics:** Game statistics are provided in the lower half of the left hand side of the screen. The statistics include: the number of bricks that have been used as well as the total number allocated for the challenge, the names of both players, the view the partner is in, the area/volume target (for SGM), and the most recent calculation of area/volume.

- **Feedback box:** Below the main building area is a black feedback box which displays information and error messages to the player. For example, if a resize request calls for more than the available number of bricks, the player is informed by the message "Not enough bricks for resize operation". Players are also informed if their partner enters 3D-View or is waiting to end the challenge.

- **Button bar:** In addition to the statistics, which are constantly displayed, two buttons on the right-hand button-bar also provide information to the player. The "i" button gives information on the current selected object (e.g., for a wall it gives the length, height and orientation of the wall). The large button marked "Area" [4] calculates the current area of the house. The algorithm to calculate floor-area works by starting at a grid which is "outside" the house (for this purpose an unseen row and column of grids is added so that the program always has an outside grid to start at). It then marks all grids it can "touch", including diagonally-touching grids, as indicated by the walls in Figure 4.2, which do not enclose an area. [5] After this the algorithm sweeps the space in search of an unmarked grid. Upon finding one it uses the same process to mark all "touching" grids, then tallies up this group of grids to get the area for that "room". This sweep-and-tally process is continued until the sweep finds no unmarked grids. The algorithm terminates and feedback to the player is provided in the following manner. Each of the rooms is painted a different colour and the area value is displayed in this painted region. The total area of the house is displayed in the feedback box at the bottom of the screen. [6] If the current arrangement of walls does not enclose a group of grids, no unmarked grids will be found in the first sweep, and the value returned will be zero. Any rooms with an area greater than 100 square units are painted black and the message "TOO BIG" is displayed in the room. The

[4] For the last two challenges it is marked "Volume".

[5] Apologies for the pic, it was done in LaTeX :-)

[6] The total is displayed on both players' screens but the individual room feedback is provided only on the screen of the player who clicked the button.

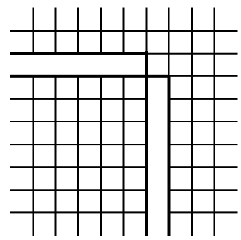

Figure 4.2: Corner of a room that does *not* enclose an area

area of such rooms does not contribute to the total area. Volume calculations take the height of the lowest contributing wall for each room and multiply it by the floor-area.

View Control
From Top-View players can move to one of two other views of the house. Moving to another view does not affect the partner's view. Under the shadow-box are the two buttons that toggle the player between Top-View and Side-View. The current view is indicated by the depression and highlighting of the appropriate button. Players can move to the 3D-View by clicking the house button on the right-hand button bar. The other buttons relating to the 3D-View (the eye and light buttons) are discussed in the 3D-View section.

Game Control
Players can also move to either the Challenge-Selection or Challenge-Info screen from Top-View. The question-mark, as previously mentioned, returns the player to the Challenge-Info screen without ending the challenge. To exit the challenge, and hence move back to the Challenge-Selection screen, players click the red "Q" button. If Dave clicks this button while Sonia is still building, Dave will be sent to a waiting screen, from where he can either wait for Sonia to click "Q", or choose to re-enter the challenge. If Sonia clicks the "Q" while Dave is at the waiting screen, both players are returned to the Challenge-Selection screen. If they have successfully completed the challenge, the appropriate button will be coloured yellow to indicate competition and their score for the challenge will be displayed.

4.4.4 *"Side-View" (Screenshot 4)*

In Side-View players can add windows or doors to the existing walls of their house. Adding windows and doors contributes to the successful completion of the challenge by freeing up bricks which players can then use for making more walls or enlarging existing ones. If the player already has a wall selected when s/he enters Side-View this wall is displayed from the side. No other walls are visible. Below the wall (on the "grass") are data on the current surface area and bricks consumed by the active wall. The player uses the door and window buttons on the right-hand button-bar to create the objects. The wall button is inactive in Side-View, and will produce an alert sound if clicked. Newly-created windows and doors appear in a default location, as for walls, but are different in that they are initially incomplete, which is indicated by dotted lines. To complete construction of doors and windows players must choose their horizontal and vertical frame pieces from the limited store provided at the beginning of each challenge. A door or window, which we will refer to as an "object", does not become included in the active wall until both of the following conditions are met:

- both horizontal and vertical sets of frames of the object have been set;
- the object is entirely contained by the active wall, and does not overlap any existing objects.

The various possible states are indicated by the colour of the object and its outline. When first created the whole outline of the object is dotted. Initially, since the object is active, the dotted outline is coloured yellow and black. If the object is not active (i.e., not the currently-selected object) the outline is coloured white and black. When the player chooses a set of frames the appropriate two sides of the object [7] go from dotted to solid lines (yellow if active, black if not). The top two buttons on the right-most column of the button-bar set the horizontal and vertical frames respectively, and when clicked present a dialog box listing the frames still available and asking the player to type the length of the desired piece. Until the object becomes officially part of the wall it is coloured pink, and in Top-View will not be visible. Once the player has placed the object so that it is fully contained by the wall and not overlapping any existing objects, the image of the object changes from pink to a picture of a door or window. At this point the server adds the object to the wall, which results in the bricks covered by the object being returned (the player is notified by sound and visual feedback, and told exactly how many bricks have been returned), and the object appearing on the appropriate wall in Top-View.

A window or door can also be removed from a wall, at which point the object will again become pink, the released bricks taken back, and the object will disappear from Top-View. An object cannot be removed, however, if there are not enough bricks available to fill the hole it would leave. Pink objects remain in the current Side-View even when the player switches between walls. In this way the player can choose to place an already-completed window on whichever wall s/he wishes. The player can switch between frame sizes at any time, or delete the whole object, which returns the object's frame pieces to the store.

4.4.5 *"3D-View" (Screenshot 5)*

The 3D-View allows players to move around and inspect the house that they have constructed. There are three buttons on the button bar in Top-View relevant for the 3D-View (shown in Screenshot 3): an eye button which allows the player to set the viewing location and direction; a light button by which players can place an additional point-light source in the scene; and the house button which sends the player into 3D-View. Once in 3D-View the player can navigate around using the arrow keys. Walls appear in red with holes where windows and doors have been placed. Players can move through objects and so can enter into the house and see what it looks like from the inside. In ECM, if Sonia is simultaneously

[7] That is, the top and bottom if the player has chosen the horizontal frames, or the left and right if the player has chosen the vertical frames.

in 3D-View and in Dave's field of view, Dave can see Sonia's avatar moving around and vice-versa.

4.5 The Client-Server Structure

As *Builder* is a networked multiplayer game, distributed programming is necessary, meaning more than one process runs simultaneously and processes must communicate with each other. This is done using a client-server model, which this section briefly describes. In a client-server model clients do not communicate with each other directly, but rather communicate with a central server. *Builder* was originally implemented using a client-client model with clients communicating directly with each other, but was converted to the client-server model after the first pilot study (Section 6.1) to address synchronization and other stability problems.

Builder runs within the *Island* framework, with the servers and clients implemented as C++ classes. There is a base class `Server` and a base class `Client`, and `IslandServer` and `IslandClient` (hereafter IS and IC) and `BuilderServer` and `BuilderClient` (BS and BC) are all children (at the same level) of the two base classes. All other activities within *Island* are implemented in the same way as *Builder*. IS and IC are always active, while each of the other server-client pairs is active only when the player is engaged in the specific activity. Whenever an IC or BC is created, it checks whether there is already a server running, and if not it creates one. Other machines are notified of where each of the servers are running so that they know where to send messages. These network messages are received initially by the IC on the machine running the server in question and passed on to BC if appropriate. The servers do not run as separate processes but are rather collections of functions that are called by the client on the same machine. IC contains the main program loop. While *Builder* is running, most user events are handed off (by IC) to BC. Exceptions to this include events relating to outgoing written and spoken messages, which are handled by IC. In addition to user events, IC also handles all incoming messages arriving over the network. Incoming written or spoken messages are dealt with directly by IC (displayed to the message window or played through the speaker). All *Builder*-related messages are sent on to BC. As BC processes the user events it determines when calls to BS are required. Anything that alters the collective game information or the user's status in the game requires a message to the server. For more information on messages passed between BC and BS see Appendix B.

Chapter 5

Tools for Assessing Outcomes

The study forming the central part of this thesis compared the effect of various conditions on a set of outcomes. The independent variables of the study were: mode of communication (written vs. enhanced), nature of task (specific vs. general), and gender(male vs. female), and will be discussed further in a subsequent design section (Section 7.1). This section is concerned with the dependent variables, and the tools developed to measure them. The main outcomes of interest were: academic achievement, performance in the game, the nature of game-play, and attitudes toward the task and the playing partner. Formal tools were developed to assess academic achievement and attitudinal outcomes. Performance in the game and nature of game-play were assessed via automatically-recorded game logs and anecdotal observations recorded by the researchers.

5.1 Academic Measures

5.1.1 *Target Areas*

The target learning areas, as discussed under the section on Task above (Section 4.2), were:

- the relationship between perimeter and area;
- addition and subtraction of areas and volumes;
- tiling of surfaces.

5.1.2 *Tests*

Academic improvement in the three target areas was assessed using a pre-test and a post-test, which are presented in Appendix C. The tests each consisted of 10 items, with each of the post-test items being a variant of the corresponding item in the pre-test. The original versions of the tests were developed by the researcher based on inspection of standardized mathematical test materials [CToBS88] and textbooks [Les86, e.g.] for the Grade 7 level. The tests were then revised based on pilot testing (see Pilot Study 1, Section 6.1) and consultation with researchers in mathematics education. The final structure adopted was as follows.

- Three items asked for simple calculations concerning area, volume, perimeter and tiling.
- Three items asked for calculations regarding particular scenarios from the activity – e.g., calculation of floor area given the length of walls, calculation of bricks saved by windows of a certain size, etc.
- Four word problems required understanding of the target concepts but in different contexts – e.g., how much area has a grass-cutter of width X cut after one cycle around a park of dimensions YxZ.

The only difference between the pre-test and post-test were the numbers involved in the calculations, and the scenarios used in the word problems. The comparability of the two tests, and the assumption that the items on each test assess a unitary underlying construct, [1] are addressed at the beginning of the results section.

5.2 Socio-motivational Measures

Appendix D presents the questionnaire that was completed by subjects directly after playing the game. Twenty attitudinal items were included to assess the sociomotivational outcomes of game-play. For each item, students indicated the degree to which they agreed or disagreed on a 5-point, Likert scale (YES, yes, maybe, no, NO). The dependent variables which these items were designed to assess were as follows:

- attitude toward the game (e.g., "I enjoyed playing *Builder*", "I learned something by playing *Builder*");
- attitude toward partner (e.g., "If I play *Builder* again I would like to play with the same partner", "My partner was friendly");
- desire to continue playing (motivation) (e.g., "I would like to play *Builder* again", "I would like to play *Builder* at home");
- perception of collaboration (e.g., "I would prefer to have my partner in the same room", "Communicating with my partner helped us to play the game").

In addition to the attitudinal items, there were three background questions regarding home computer use. These questions were used to gather descriptive information on patterns of use amongst the sample pool, and to ascertain whether these patterns reflected the gender differences outlined in the literature review.

Items on an initial version of the questionnaire were modified on the basis of pilot studies and consultation with a psychology research group, the latter of which suggested adding related items for each of the target dependent variables to improve the psychometric

[1] It is possible, especially given the three categories of question discussed above, that different items draw on different knowledge or achievement domains. In this case it would be inappropriate to use only the whole test scores in data analysis, as we could be averaging over important differences that would be seen if the test items were broken into domain-specific groups.

robustness of the questionnaire. The research group also identified the need to obtain a pre-game measurement of partners' liking for each other, without which it would be dubious to make comparisons between conditions based only on their responses to the post-game questionnaire. This was addressed by introducing a one-item pre-game questionnaire which asked subjects to answer the question: "How much do you like playing with your allocated partner (<name-of-partner>)?" on a 5-point scale between "Not at all" and "A lot".

5.3 Log Files

Builder logs two types of information for each session of play. First, for all completed challenges, a record of the challenge number and attained score is kept. As mentioned in the chapter on *Builder*, the challenge score indicates either the number of bricks used (for SGM) or the area/volume attained (for MGM). Scores are only recorded if challenges are successfully completed, defined either as achieving the target (SGM) or enclosing a non-zero area/volume (MGM). The number of challenges completed and the challenge scores were used as measures of performance in the activity.

Second, information about the messages passed between players is recorded, including a script of the entire written dialog. We decided not to record speech messages because of the overhead of writing such large chunks of data to file during play. Instead, at the completion of play, a measurement of the total amount of speech data sent over the network is written to the log file. These logs of written and spoken communication allow for comparison of how, and how much, subjects in different conditions communicated.

5.4 Observations

During each session of game-play, researchers also completed observation forms (shown in Appendix E) to supplement the information in the log files. As the forms indicate, one of the main issues here was the type of communication between players. This was particularly important for players using spoken communication, since the log files did not record spoken messages. The forms also allow for ad hoc observations which are often useful as suggestions for further research or game improvements. For example, if a player repeatedly attempted to resize or move an object in a way that the interface does not support, this could be noted in the observations and considered as a modification to the interface.

Chapter 6

Pilot Studies

Four pilot studies were conducted prior to the commencement of the final study. Such preliminary field testing was necessary given that both the software and the assessment tools were entirely new. The pilot studies also played an important role in the design of the interface and other aspects of the software. The dates and locations were as follows:

- **Pilot Study 1:**
  - *Date:* March 11-13, 1997
  - *Location:* Trafalgar Elementary, Vancouver
  - *Subjects:* 24 grade 6/7 students
- **Pilot Study 2:**
  - *Date:* April 24, 1997
  - *Location:* E-GEMS laboratory, UBC (students visiting from Island Pacific School, Bowen Island)
  - *Subjects:* 8 grade 7-9 students
- **Pilot Study 3:**
  - *Date:* May 7, 1997
  - *Location:* Trafalgar Elementary, Vancouver
  - *Subjects:* 14 grade 6/7 students
- **Pilot Study 4:**
  - *Date:* May 20, 1997
  - *Location:* Kerrisdale Elementary, Vancouver
  - *Subjects:* 12 grade 5 students

For each of the pilot studies, the duration of one session of play was approximately 30 minutes. As was the case for the final study, the number of subjects for these studies was limited by the number of Power Macintoshes available for play. With the exception of Pilot Study 2, playing was limited to the two machines brought into the school by the researchers, allowing only one pair of students to play at a time.

6.1 Description of Pilot Study 1

The initial pilot was the most extensive in terms of duration and size of subject pool, and had the most influence on design changes because it was the first opportunity to realistically evaluate the user interfaces and stability of the software. Therefore it will be discussed in detail.

6.1.1 *Goals*

The initial pilot study had four objectives. The primary objective was to evaluate the usability of the game with respect to the following questions.

- Does the software behave reliably under "real" conditions; i.e., when played by two people of the target age-group for a reasonable length of time?
- Is the game interface understandable to first-time users?
- Are the challenges within the game of an appropriate level of difficulty for the target age group, given the intended duration of play in the final study?
- Is the game sufficiently stimulating to engage players for the duration of play?

A secondary objective was to evaluate the effectiveness of the proposed assessment methods. This was particularly important for the pre- and post-tests assessing academic outcomes, where it was necessary to determine whether they were of an appropriate level of difficulty for students in the target age group, and whether the pre- and post-test were comparable. Our goal for the level of difficulty was for mean test results of approximately 50% on each test (assuming that there is no game-play or other form of task-related instruction between the two tests), thereby allowing a wide spread of marks and a sufficient margin to observe improvement on post-test scores. For the questionnaire, it was also necessary to ensure that all items were readily understood by the students.

A third objective was to evaluate the proposed set of independent variables for the final study. The main independent variable being considered in Pilot 1 was mode of communication (written vs. spoken). Accordingly, subjects were split into two groups: those with both written and speech communication, and those with only written communication. [1] Furthermore, as planned for the final study, subjects were tested in same-sex pairs so that there was an opportunity to observe gender differences in playing.

Finally, the pilot allowed the researchers to practice running the study. A study of this nature involves juggling a multitude of practical constraints, and pilot studies are very useful in identifying problems in the procedure that might compromise the final study.

6.1.2 *Procedure*

Prior to the first day of Pilot 1, consent forms were distributed to students in order to obtain parental permission for their involvement in the study. Only those students who received parental permission and who themselves agreed to participate were included in

[1] The virtual presence element had not been implemented at the time of Pilot 1.

the sample. On the first morning of the study, two Power Macintoshes were set up in the library of the school so that there was no direct visual or aural contact possible between the players at the two computers. The researcher then gave a 10-minute orientation to the class, introducing the students to the task and interface of the activity. The orientation was followed by the pre-test (completed in the classroom), for which students were allowed 20 minutes. Following completion of the pre-test, the dyadic playing sessions began, with same-sex pairs being taken in turns to the library area where the computers were set up. Each session began with a researcher briefly explaining the interface to the student at each computer. This consisted simply of pointing to some of the main buttons and briefly describing their function. Although a help screen was provided, few players spent much time reading it. The pair then played the game for approximately 30 minutes. When their time was up, the students were seated together in another area of the library and asked to complete both the post-test and the attitudinal questionnaire, for which they were allowed as much time as they needed. In addition to the 16 5-point scale attitude items, [2] the questionnaire included the following open-ended items that allow students to identify problems with and propose enhancements to the game.

- Please write anything you found frustrating about playing Builder.
- Do you have any comments or suggestions on the communication?
- Do you have any other comments or suggestions about Builder?

6.1.3 *Results of Software Testing*

Performance of game

The observations of the game's performance in Pilot 1 revealed the presence of several bugs in the implementation. At the time of Pilot 1, *Builder* used a client-client distributed model where clients inform each other of their actions, rather than informing a central server. Although a simple form of locking objects and actions had been implemented, observations indicated that delays in the receipt of between-client messages resulted in inconsistent information about the state of the game, and occasionally in one of the machines crashing. [3] Crashes impeded game-play considerably because players had to wait for the machines to reboot, which took several minutes.

Changes made to improve performance

While these problems could have been fixed within the client-client model by implementing a thorough verification system, we decided to reorganize the code to follow the conceptually simpler client-server model, in which the server enforces consistency of information. These changes took several weeks as it was also necessary to restructure the central *Island* code, and, because the changes were quite extensive, led to the need for further field testing.

[2] The 5-point items contained in the original questionnaire are similar to those of the final questionnaire (presented in Appendix D, items 1–20).

[3] Crashes were mostly the result of one program trying to act on messages from the other program that referenced an object about which the machines held inconsistent information.

6.1.4 *Assessment of the Interface*

The following assessments regarding the design of the game are drawn from the observation forms and the students' feedback on the open-ended items in the questionnaire. Many of the points relate to "dressing up" the game's interface. The changes that were made are also described. Most of these were made to improve the ease-of-use and/or enjoyment of the game, and relate primarily to providing adequate feedback to the users.

Feedback

- Players did not appear to be sufficiently aware of what their partner was doing. This was partly related to the degree of communication between partners which is discussed further below. At the time of Pilot 1, the wall currently being used by the partner was indicated only by a blue highlight around the small image of the wall in the shadow box. [4] It appeared that most players did not pay attention to this. This was evidenced, for example, by one wonderful response on the questionnaire: "The computer started moving and adding walls." Therefore the interface was modified so that the partner's wall was outlined by a blue highlight in Top-View as well as in the shadow box. Further pilots indicated that players were still not noticing the blue highlight, so in the final version the partner's wall was given an entirely different colour, rather than just a highlight.

- In addition to the problems of awareness of the partner's actions, there was evidence of a need for more feedback in general. While some illegal actions were flagged by an alert box, [5] for many important actions there was no feedback. Rather than adding alert boxes for all these actions, we decided to add a feedback box to the bottom of the main building area, which would display messages regarding important actions to the player, and remove the interruption of play caused by alert boxes. Messages to be displayed in the new feedback box included confirmations of the player's actions (e.g. "Moved wall"), notification of the partner's actions (e.g. "Partner is waiting to end the challenge"), error messages if the player's action is illegal (e.g. "Partner's wall" when the player tries to select the partner's current wall), and game information (e.g. "Area: 90", after the player clicks on the area button).

- Feedback for area or volume attained was also improved as a result of Pilot 1. At the time of the pilot the feedback consisted simply of displaying the total area or volume of the house as a number on the screen. We decided to enhance feedback in this area by calculating the area/volume of each room separately, and then painting each intact room a different colour and displaying the area/volume in the middle of the room itself (while displaying the total in the feedback box). Furthermore, to draw attention to the problem of exceeding the room-size limit, we decided to paint oversized rooms black and mark them with the message "Too big!".

Enjoyment

[4] The shadow box is the miniature copy of the top view, as described in Section 4.4.3.

[5] An alert box, typical in Macintosh and Windows applications, pops up within an application like a dialog box to present important information, and must be clicked by the user before work within the application can continue.

- Based on suggestions from several subjects on the questionnaire, and also for the sake of additional feedback, it was decided to add sound effects to *Builder*. Sounds added include: generic clicking sounds for buttons, specific sounds for flipping, moving and resizing objects, and various different alert sounds.

- Two of the sound effects added addressed specific feedback issues. First, researchers identified the need for feedback regarding bricks being returned after the placement of doors or windows. This is an important concept in the game because of the connection between the amount of bricks released and the lengths of the frames which the player has selected. There should also be a sense of reward attached to the procurement of more bricks to build walls with. Therefore, to draw greater attention to the bricks returned, we decided to add a sound effect emulating a slot machine and to display the exact number of bricks returned in the feedback box. Second, an alert sound was added to draw attention to incoming messages, which players often failed to notice.

- Among the responses were several specific enhancement suggestions for the 3-D component of the game, which appeared to generate a substantial amount of interest. Most common among these was the request for animation, which had in fact been partially implemented at the time of Pilot 1, but was not ready for inclusion in the version of *Builder* used in the study. Along with the facility to move around the 3-D space we added moving avatars to represent players when both are in the 3D-View simultaneously. The effect of this addition was included in the design of the final study (in the mode of communication comparison).

Ease-of-use

- One respondent requested the addition of a button to send messages rather than having to use the menu bar, which was implemented in the revised version.

- The interface for moving walls in *Builder* is drag-and-drop. In the initial version, the first click on a wall initiated the drag-and-drop process. Many players showed some frustration when they clicked on a wall with the intention of selecting it (e.g. for resizing or flipping) and the wall then jumped to another grid, because they were moving the mouse around while clicking. To remove this frustration, and also to provide a convenient way to implement locking, the interface was changed so that an initial click on a wall indicates only selection. The wall must then be clicked again to initiate the drag-and-drop. [6]

- The controls for switching between Top-View and Side-View were not prominent enough. Initially implemented just as words on the screen that players had to click, they were updated to look like buttons. All buttons in the interface were modified so that they appeared raised out of the screen, and depressed into the screen when clicked. Inactive buttons were either hidden or darkened.

[6] There is often a trade-off, however, with such decisions. In the updated version of the game players were frustrated when they clicked on a wall to initiate the drag-and-drop and the wall did not move. It might be possible to refine the interface further by measuring how long the mouse button is held down and hence distinguishing between clicks intended for selection and grabbing.

- Several players counted grids on the screen to try to calculate the interior lengths of walls, which appeared inconvenient. One respondent also complained that the grids were too small for counting. On the basis of this, we added an information button (marked "i") to provide dimensions of objects when clicked, though many subsequent players nevertheless continued to count grids. Although this might change if players had had a longer time to familiarize themselves with the interface, counting grids is a useful activity in terms of reinforcing the players' understanding of length and area.

- The need to type sizes into dialog boxes for resizing objects was observed to be unintuitive for players. When presented with a list of sizes to choose from and asked to type one of them (which was the case for choosing frame pieces) players often tried to click on items in the list rather than typing. The frustration seen with the dialog-box method of resizing raised the design question: *what is the best interface for resizing?* To address this question an alternative resize interface was implemented, which is described below in connection to Pilot 3.

Other areas addressed

- Modifications were made to the area and volume goals for some of the challenges to adjust the level of difficulty. A more important change which was incorporated later was to require that the area/volume goals be met exactly. At the time of Pilot 1 the goal was to *at least* reach the stated area/volume target. It was thought that making the goal more rigid would draw greater attention to the numbers involved and also make the challenges more difficult to achieve by trial-and-error. These conjectures were strongly supported by observations in the final study, which indicated that many players did not start to think about the lengths of their walls until they had had the experience of making the house both too big and too small.

- The statistic of the overall surface area of the house was removed because it was rarely looked at, and when it was, seemed to be confusing rather than helpful.

- Following observations that players spent negligible time in the help screen, it was removed and replaced with the Challenge-Info screen (as described in Section 4.4.2).

6.1.5 *Findings Regarding Communication and Gender*

Statistical analyses were not performed on the results of the pilot, since the interruptions caused by the instability of the software may have had a disproportionate effect on different groups. There were, however, several interesting communication- and gender-related game-play observations. To begin with, there was evidence of a wide variety of different collaborative strategies. The most highly-structured approaches were characterized by partners telling each other what to do, informing each other when they had finished a task, requesting confirmation of what the other was currently doing, giving reasons for their actions (e.g. "I'm trying to save bricks") and even initiating turn-taking systems. At the other end of the spectrum were pairs that used the communication mainly for fun or for insults. In the latter case, partners would sometimes not communicate at all until the partner had done something annoying. There was a final category of players who barely communicated at all, and this did not appear to be related to whether they had speech communication available to them or not.

Some of the observed communication differences appeared to interact with gender. For example, it was noted that some players with access to both forms of communication still showed considerable use of the slower written medium, and this was more often the case for girls than boys (one girl even used the spoken communication to tell her partner to write to her more). Most boys with speech tended not to use written messages much. Another notable gender difference was that some girls seemed more focussed on the communication than the task, which was almost never true for the boys.

6.1.6 *Difficulty and Engagement*

The degree of difficulty appeared to be appropriate for the pilot sample. Most players were able to successfully complete at least one of the challenges, suggesting that the activity was not prohibitively difficult. Some players completed two or more of the challenges, but none completed more than three, indicating that the game was not too easy. As mentioned above, the difficulty of some challenges was slightly modified by changing the area/volume goal or the number of available bricks.

The findings regarding degree of engagement in the activity were encouraging. In contrast to the researchers, players seemed relatively unconcerned by any problems encountered, and many expressed a desire to play again. Most of the comments on the questionnaire expressed a positive attitude toward the game. Several players said the game was fun to play, and about the same number again specifically mentioned that communicating was fun (referring both to written and spoken communication, but especially the latter).

6.1.7 *Findings Regarding the Test Materials*

The pre-test produced a satisfactory range of results, with a mean of just less than 50% (about 2.2 out of a possible 5). The teacher of the class confirmed that the results reliably predicted known math level. Post-test performance, however, was lower than expected, and, surprisingly, lower than pre-test scores, with an average of around 1 out of 5. The low average was an obvious concern, though we did not interpret this as implying that playing the game was detrimental to math performance. Rather, we concluded that items on the post-test may have simply been more difficult. Accordingly, pre- and post-tests were restructured so as to be more symmetrical. The pilot was helpful in this restructuring since it gave a good indication of the difficulty of each item. The low post-test scores may also be attributable to the administration of the post-test, which is discussed in the following section. Given that the results of the pilot study were open to the interpretation that playing the game had a negative effect on test performance, it was decided to add a control group for the final study. Control subjects would sit both pre- and post-tests, and the difference between their scores would provide a baseline against which to compare the test results of the experimental groups.

There were several minor changes made to the questionnaire after Pilot 1. It was found that one of the items was confusing because it was expressed in the negative and therefore hard to answer on the yes-no scale. This item was reworded. The open-ended section was removed for the final study because it would increase the time required for each subject, and because the questions regarding problems of and improvements to the game were specifically relevant during the development of *Builder*.

6.1.8 *Procedure-related Findings*

The main procedural concern that emerged during Pilot 1 was in regard to the administration of the post-test and questionnaire. There was some evidence of partners making identical responses to the questionnaire, and it also appeared as though relatively little effort was put into the post-test, compared to the pre-test. For the pre-test, the whole class was present together with the teacher supervising, as for any normal classroom test. Therefore there was some pressure to perform well on the test. Furthermore, since the teacher maintained silence for the allocated time limit, students were encouraged to spend more time thinking about the questions. For the post-test, however, students were seated together and left unattended, allowing them to discuss responses and also removing the aspect of test pressure. Students also knew they could leave whenever they were done, which may have provided the temptation to skip over time-consuming questions which the classroom setting would have prompted them to attempt.

To address this concern, two procedural changes were introduced. The first was to delay the administration of the post-test until after all participants had played *Builder*, and thus allow the test to be taken as a class in the normal classroom setting. This had the disadvantage that the game would no longer be fresh in the subjects' minds, but had the advantage that the test would be taken more seriously and performance would therefore be comparable with that on the pre-test. The second change was to separate the partners while writing the questionnaire (which would still be done directly after the game – since it was not mentally taxing like the post-test, there would be less reason for students to skip items). It was hoped that this modification would encourage students to provide considered responses based on their own experience.

The other major change to the procedure was in regard to the specific pre-game orientation. First, so that the orientation would be standard across all participants, it was formalized and always given by the same researcher with both members of the dyad in front of the computer (which had the additional advantage of freeing one of the researchers to oversee the last pair still working on the questionnaire). During the formalization of this orientation, several important features were added (see Section 7.4). One addition that emerged specifically from Pilot 1 was the decision to explicitly encourage players to communicate with each other during game-play. A lack of effective communication between partners was frequently observed during the pilot study, though there were certainly exceptions to this. The extent to which players communicated was entirely up to them, with the intention being that if they felt a need to communicate they would do so. This approach was reconsidered in light of the observations, as well as illuminating student responses such as: "Make players tell each other what they're doing more often". It was also thought that explicitly encouraging players to communicate was in accordance with the afore-mentioned Cooperative Learning technique of training learners in collaborative techniques. In addition to the human communication problems, there were also technical problems such as fuzzy and choppy messages. These were addressed by including, in the description of how to communicate, reminders to hold the Control button right until the end of the spoken message and to not talk directly into the microphone.

6.2 Other Pilot Studies

Due to the extensive restructuring of the *Island* and *Builder* code discussed above, further field testing was essential to ensure that the software was sufficiently reliable for use in the final study. The other pilot studies, however, will not be discussed in detail, since the problems identified and the resulting changes were considerably smaller, both in number and importance, than those from Pilot 1. Pilots 2 and 3 revealed the existence of problems with the modified software, including a serious memory leak which occasionally led to crashes. By Pilot 4 the memory leak had been located, and the number of problems had been reduced to the extent that only a single bug was identified during the six sessions of play.

Between Pilot 1 and Pilot 2 the pre- and post-tests were extensively modified, both approximately doubling in length. This necessitated further testing to again attempt to balance the two tests for level of difficulty. Hence the academic tests were administered to students in both Pilot 2 and 3, and there was some resultant shuffling of questions.

6.2.1 The "Unintuitive" Interface (Pilot 3)

As mentioned above with regard to interface problems identified in Pilot 1, there was some concern over how best to allow players to set the size of walls, windows and doors in *Builder*. Observations indicated that typing a number into a dialog box to set the size of an object was "unintuitive", where we define "intuitive" as being the way that naive users naturally attempt to perform an operation. As an experiment in interface design, we decided to implement an alternative method of setting the size of an object. The alternative interface, which we will refer to as *drag*, does away with the dialog box, and instead displays a small green dot at the bottom-right corner of the object, which users can drag to reset the object's size. As the dot is dragged (i.e., while the mouse button is still depressed), feedback is provided dynamically in two forms. First, a "shadow" rectangle appears, whose top-left corner is set at the top-left corner of the object, but whose bottom-right corner follows the moving cursor. [7] Second, the dimensions of the object at the current position of the dot are displayed to the user via the feedback box. In Top-View, for example, if the dot is dragged such that the shadow rectangle is ten units long, the feedback box would say "Current length: 10". (Note that in Top-View the only possible size modification is to the length of a wall, so when players drag the green dot the shadow rectangle changes only in length, not in width, since the width of a wall is set at one unit.) In Side-View the situation is more complex, since feedback is required on the current width and height of the window or door. Furthermore, the program dynamically calculates the nearest available horizontal and vertical frame lengths corresponding to the current cursor position, and displays these two numbers in the feedback box. Upon release of the mouse button, providing the new size is legal, walls are enlarged or reduced to the nearest grid-line, and windows and doors are set to the nearest available frame lengths.

[7] This "shadow" rectangle is familiar to anyone who has dragged the bottom-right corner of a window in most current Macintosh, Windows, and even UNIX-based X-Windows applications.

The *drag* interface would be judged a "better" or more "intuitive" interface by most HCI criteria. It involves direct manipulation, giving the illusion of active stretching or shrinking of an object. It also provides interactive feedback, with the shadow providing a WYSIWYG-style preview of what the modified object will look like. In the *dialog* interface, on the other hand, naïve users may have no idea of what the result of their action might be, and may therefore be tentative about typing in a new number. However, returning to the discussion of HCI issues in the literature review, it is not clear whether an intuitive interface is appropriate within a game or educational application. Within the *Builder* context, where the aim is for learners to experiment with mathematical concepts, it is important that attention is paid to the object sizes. An intuitive and easy-to-use interface like *drag* has two potential disadvantages. First, because users do not have to think of a number to type into the dialog box, they may be less aware of the size they have set the object to. Second, because it is faster and easier than *dialog*, users are more likely to use trial-and-error, continually modifying objects without reflecting on what they are doing. As a result of the preceding points, when players in SGM, for example, reach their area/volume goal, they may not be aware of what size the walls have been set to, which makes it less likely that they will make any discoveries about interior area in relation to side-length.

The *drag* interface was pilot-tested during Pilot 3, where half the pairs used *drag* and the other half used *dialog*. At this point in the development of *Builder*, the interface issue was being considered as an additional independent variable to manipulate in the study. The hypotheses for the different interface styles were:

- *drag* will be considered easier and more enjoyable than *dialog*
- *dialog* will result in better performance on the post-test

Unfortunately, it was not possible to look at pre-post comparisons for Pilot 3, due to the fact that the tests were still being calibrated and because some of the participants that played *Builder* during Pilot 3 had also played in Pilot 1. Interviews with the players who had used both styles, however, confirmed that *drag* was preferred as an interface style. Although indications suggested that the interface comparison would make an interesting variable for the final study, it was decided not to include it because it does not bear directly on collaboration in the game. Therefore Pilot 3 is best viewed as the preliminary investigation of a possible future study. Based on the potential advantages of the "unintuitive" interface discussed above, *dialog* was chosen as the interface for the final study.

Chapter 7

Study Design and Methodology

7.1 Design

The experiment was a 2x2x2 factorial design. The **independent variables** were as follows:

- **gender (SEX)** – male-male pairs vs. female-female pairs
- **mode of communication (COMM)** – basic communication mode (BCM) vs. enhanced communication mode (ECM)
- **nature of task (GOAL)** – specific goal mode (SGM) vs. maximize goal mode (MGM).

The **dependent variables** were:

- **academic gain** – post-test score minus pre-test score
- **game performance** – number of challenges completed and challenge scores
- **sociomotivational effects** – attitudes toward the activity, attitudes toward partner, persistence of interest in game, and perceived collaboration.

A no-activity control group was used to provide a baseline for pre- and post-test comparisons, so that any differences seen in the experimental conditions could be judged relative to differences seen in the control group. It was hypothesized that the game group (composed of all students who played the game, regardless of GOAL or COMM condition) would show greater academic gain than the control group. All comparisons other than this initial game vs. control test were between the eight cells of the 2x2x2 design. Based on theories and findings from Cooperative Learning, it was hypothesized that the ECM group would show higher achievement, both in terms of academic improvement and game performance, than the BCM group. For the nature of task hypothesis, the situation was less clear. Based on findings that ill-structured tasks lead to more effective Cooperative Learning [Coh94], it might be expected that MGM dyads would show greater learning improvement than SGM dyads. However, a specific goal may make players focus more on the numbers involved (especially relevant for mathematical learning), and may also be more motivating due to the direct feedback of knowing whether or not the goal has been attained. Hence no clear hypothesis was made on the effect of nature of task. [1] For the socio-motivational outcomes, it was expected that ECM would lead to more positive attitudes across each of the four categories identified above. No specific hypotheses were made for gender, since the intent was to discover differences that should be addressed in CSCL design, rather than to support an explicit theory. On the basis of the computer game research discussed in the literature review, however, we might expect males to perform better on the task. Given the goal-oriented nature of the game, we might also expect that males would be more positive than females towards *Builder*. On the other hand, the fact that the game centers around communicating and working together might lead us to expect a more positive response from females.

7.2 Participants

The participants of the study were 134 students from two elementary schools: Queen Elizabeth Elementary in Vancouver and Diefenbaker Elementary in Richmond, and had received parental permission for their participation in the study. All participants were from grades 6 and 7 (10-12 years old). There were 100 students who played the game, 48 girls and 52 boys. The other 34 students were in the control group and therefore completed only the pre- and post-tests. To minimize the confound of sampling from two different populations it was ensured that half of the subjects in each of the cells in the study design came from each school. Prior to the commencement of the study, permission forms which explained the purposes of the study were sent home with the students to obtain parental consent for their participation.

7.3 Materials

- Hardware: two Power Macintosh computers connected via AppleTalk.
- Software: the game *Builder*.
- Tests: written pre-tests and post-tests containing mathematical questions related to the concepts in *Builder*, as described in Chapter 5.
- Questionnaire: the written form described in Chapter 5, consisting of 20 5-point Likert scale items assessing socio-motivational outcomes of the study, and three questions on home-computer use.

7.4 Procedure

Following is the sequence of steps followed in the study, as administered at each of the schools.

[1] Regarding game performance, it was not possible to consider differences for the two GOAL conditions because of the different nature of the challenges undertaken in either condition.

- *Written mathematical pre-test*
  Several days before the first dyad from each class played *Builder*, the first of the two academic tests was administered by the teacher during normal class-time. Students were allowed 30 minutes to complete the test.

- *General orientation*
  Between the pre-test and the start of game-play, a 5-10 minute informal presentation by the researcher was given in front of the class to put the research and game in context. It was hoped that introducing some of the main features of the activity (building a house using walls, windows and doors; the various 2-D and 3-D views) would help players during the specific orientation and game-play phases. This also served to introduce the researchers to the students in the hope that students would feel more comfortable when they came in pairs to play the game.

- *Specific orientation*
  A 5-minute explanation was given by the researcher with each pair in front of the screen of one of the computers immediately before they commenced playing. The function of each of the buttons in the interface was explained and briefly demonstrated. Players were told how to communicate (different for BCM and ECM groups), and the scoring was explained (different for SGM and MGM groups). They were encouraged to complete as many of the challenges as possible, but warned that they would not be able to complete all five challenges during the allotted time. Additionally, in an attempt to enhance the perception of positive goal dependence and individual accountability, dyads were told: "Do you remember the special math test that you did a few days ago...? There will be another test later on. For both of these tests everyone gets their own individual score. But in the game [indicate screen], your score will be as a pair. Everything is done as a pair – you share bricks, you are working on the same house at the same time. Because of this it is good to communicate as much as possible so that your partner knows what you are doing."

- *Partner pre-question*
  After the specific orientation, one student was taken to the room with the other computer. Each player was then asked to rate on a 5-point scale how much they liked playing with their partner on an everyday basis. Since partners were assigned by teachers either randomly or according to other class activity constraints, this was to test for the possible confound of some cells ending up with more partners who happened to be good friends than other cells.

- *Game play*
  Dyads played the game for 30 minutes, with one student and one researcher at each computer. During game-play researchers silently made written observations on the nature of play and communication. Assistance with the interface was given if players asked a specific question or if they were having technical difficulties (for example, if a player was trying repeatedly to click on a button when s/he first needed to complete a dialog box, the researcher would explain what had to be done).

- *Questionnaire*
  At the end of the time limit players were asked to stop and were taken to a separate table to complete the attitudinal questionnaire. If possible, the two participants were seated at different tables or in different rooms. On the occasions when this was not possible, players were asked to complete the questionnaires silently. A duration of 5-10 minutes was allowed for the questionnaire.

- *Written post-test*
  This was administered on the day after the last students from the class had played the game (for some students this meant up to five days after the playing session). The administration was as described for the pre-test.

Chapter 8

Results

The discussion of the results is organized as follows:

Dependent Variables	*Results discussed*
Achievement	(i) pre- and post-test comparison of participants who played the game with the control group (*n=134*) (ii) comparison across levels of COMM, GOAL and SEX on pre- and post-test and performance in the game (*n=100*)
Sociomotivational attitudes	questionnaire data
Game-play	log files; observation forms

For simplicity, abbreviations will be used for the independent variables and their levels, as in the previous chapter. The three independent variables of the design – gender, mode of communication and nature of task – will be designated by the terms SEX, COMM and GOAL respectively. The abbreviations ECM (*enhanced communication mode*), BCM (*basic communication mode*), SGM (*specific goal mode*) and MGM (*maximize goal mode*) will be used to designate the COMM and GOAL modes.

Unit of analysis

In studies involving dyads, it is often unclear whether results should be analysed on an individual or dyad basis. Analysis conducted using the individual as the unit of analysis are preferable in that they allow for greater statistical power due the increase in sample size (i.e., the number of data points is doubled by considering each individual separately). However, such individual data may violate analytical and statistical assumptions regarding the complete independence of observations. Even though separate scores may be available for each member of the dyad, scores of a given dyad may rise or fall together, making most analyses invalid. In the present study, data obtained regarding performance in the game was, by necessity, analysed using the dyad as the unit of analysis, since players worked together on a single game and received a single score. However, for questionnaire and pre-and post-test data, obtained for each individual, decisions regarding the appropriate unit of analysis were made on the basis of procedures developed in [DG95]. Specifically, the degree of relationship between scores obtained for dyad members was evaluated using Pearson Product Moment Correlations. Following [DG95], if there was a significant correlation (*p¡0.05, 1-tailed*) observed between partner scores on any particular measure, the dyad was used as the unit of analysis, with a single score derived from the average of the two players' scores. If the correlation was found to be non-significant, the individual served as the unit of analysis, maximizing statistical power.

8.1 Achievement Outcomes

There are two types of achievement outcomes discussed in this section: learning outcomes and performance outcomes. The former refers to the data obtained through the pre- and post-tests of task-related mathematical skills, while the latter concerns performance in the *Builder* activity, in terms of number of challenges completed and challenge scores. For the pre- and post-tests, in addition to the data from participants who played the game, there were the results of the control group. Therefore the first part of the analysis of academic (pre-post) data was a comparison between the control group and all those who played the game. This was followed by comparisons on the same data across the eight groups of the SEX x COMM x GOAL design, all of whom played the game.

8.1.1 *Results on Academic Tests*

Reliability of tests
Preliminary analyses were conducted to evaluate the comparability and reliability of the tests. First, it was necessary to ensure that the pre- and post-test measured the same underlying mathematical constructs. As there was a control group who did only the mathematical tests, we were able to address this concern by looking at the correlation between their pre- and post-test scores. [1] If there is not a strong relationship between the control group's scores on the two tests, it is likely that the tests are measuring different skills and are therefore not comparable. The result of the Pearson Product moment correlation was sufficiently high to support the pre-post comparison (*r(29)=0.74, p<.01, 2-tailed*). Another concern was whether it was valid to compare total scores on the tests, rather than clusters of related items. This is important because if the items do not "hang together" it is inappropriate to use overall test scores as the dependent variable, because we may be averaging over significant differences among sub-groups of items. Internal consistency analyses were performed on the 10 items of each of the tests with the following results: *pre-test alpha coeffecient = 0.812 (n=134)*; *post-test alpha coefficient = 0.727 (n=131).* [2] An alpha of 0.6 or above is considered acceptably high for research purposes, hence it is appropriate to use total scores as the dependent variable.

"Improvement" scores
Learning results were assessed according to the difference between the pre- and post-test

[1] Some practice effect may be expected, but this does not impact on the correlation.
[2] There were some missing post-test scores due to absenteeism on the day of the post-test.

total scores, i.e., POST – PRE. This difference score will be referred to as the *improvement score*. Two sets of improvement score data were investigated with separate *analysis of variance* (ANOVA) tests. First, scores of those who played the game were compared with the control group. Second, each of the eight cells that made up the 2(SEX) x 2(GOAL) x 2(COMM) design were compared with each other, to ascertain the effect of the different conditions of game play.

Dependence of observations between partners
To determine the unit of analysis for the improvement scores, Pearson Product moment correlations were computed for improvement scores for *partner 1* and *partner 2*. As the relationship was not significant (*r(43)=0.22, ns, 1-tailed*), suggesting a fair degree of independence in scores obtained across partners, the improvement scores were analysed on an individual basis.

Game group vs. control group
An initial GAME (Play, Control) x SEX (M, F) ANOVA was performed to assess whether playing the game, irrespective of communication or task mode, led to greater academic improvement than no instruction. Sex was included in the analysis to avoid averaging over an unseen gender difference. Results indicated a significant main effect for GAME ($F(1,126)=8.36$, $p<0.01$), with the game group ($\underline{M}=1.22$, $SD=3.56$, $N=100$) showing greater improvement than the control group ($\underline{M}=-1.23$, [3] $SD=3.31$, $N=31$). There was no hypothesis that gender would have an effect, nor was any effect found.

		SGM		MGM	
		BCM	**ECM**	**BCM**	**ECM**
Male	$\underline{M}$	1.93	2.36	0.0	1.18
	SD	3.79	4.62	3.20	4.13
	N	14	11	14	14
Female	$\underline{M}$	1.71	2.43	–0.6	0.23
	SD	2.52	3.03	3.19	3.33
	N	12	14	10	11

Table 8.1: Academic improvement across three independent variables

Within-game comparisons
Table 8.1 shows the improvement score means and standard deviations for each of the eight cells in the SEX (M, F) x GOAL (SGM, MGM) x COMM (BCM, ECM) ANOVA (now excluding the control group). Inspection of the means across each of the independent variables indicated higher scores for ECM ($\underline{M}=1.58$, $SD=3.79$) than BCM ($\underline{M}=0.84$, $SD=3.31$), and

[3] The negative result for the control group indicates that, for this sample at least, the post-test was more difficult than the pre-test. Had the tests been of exactly the same degree of difficulty (as intended), the control group's mean improvement would have been zero, or positive due to practice effect.

for males ($\underline{M}$=1.32, SD=3.92) than females ($\underline{M}$=1.09, SD=3.15), but these main effects were not significant. Only the GOAL main effect was significant ($F(1,92)=8.95$, $p<0.01$), with SGM ($\underline{M}$=2.11, SD=3.45) scoring higher than MGM ($\underline{M}$=0.28, SD=3.47). There were no significant interactions.

8.1.2 *Game Performance*

Builder logs information about how dyads perform on the game by recording a score for each challenge completed. The simplest measure of game performance is the number of challenges the dyad successfully completed. This measure alone may not be a satisfactory indicator of game success, however, because successfully completed challenges differ in terms of how optimal the solution is. In fact, we might postulate that players who did fewer challenges achieved better solutions precisely because they spent relatively more time on each challenge. Scores were assigned for challenges in the following way. For MGM the score is the area or volume attained, with the goal being to maximize the score. Raw scores can be converted to percentages, based on the formula:

```
RAW SCORE / MAXIMUM POSSIBLE SCORE FOR CHALLENGE * 100
```

For SGM, a successully-completed challenge has a fixed area/volume, so the score is in terms of the amount of bricks used to achieve the fixed goal, with the goal being to achieve a low score. These raw scores can also be converted to a percentage using the formula:

```
(INITIAL BRICK COUNT - RAW SCORE) / (INITIAL BRICK COUNT - MINIMUM
POSSIBLE SCORE) * 100
```

The two scoring methods were different in terms of both the aspect of the game they related to and how much importance was placed on them. In SGM, minimizing the number of bricks used is a secondary goal, whereas in MGM maximizing area/volume is the primary goal. Therefore it was not appropriate to compare across GOAL modes, so results were analysed separately for SGM and MGM. Players in MGM in fact completed fewer challenges overall than players in SGM (2.04 vs. 2.38).

		BCM	ECM
Male	$\underline{M}$	3.29	2.67
	SD	0.76	0.52
	N (pairs)	7	6
Female	$\underline{M}$	1.83	1.71
	SD	0.75	1.25
	N (pairs)	6	7

Table 8.2: Number of challenges completed across SEX and COMM for SGM players

A series of two ANOVAs (2x2) were performed, one for SGM and one for MGM means, as presented in Table 8.2 and Table 8.3, to compare the performance across SEX and COMM variables. Results for the SGM group (Table 8.2) indicated a significant main effect for SEX

		BCM	ECM
Male	*M*	2.29	2.29
	SD	0.76	0.76
	N (pairs)	7	7
Female	*M*	1.5	2.0
	SD	0.84	0.00
	N (pairs)	6	6

Table 8.3: Number of challenges completed across SEX and COMM for MGM players

($F(1,22)=12.07$, $p<0.01$), with males ($M=3.00$, $SD=0.71$) completing more challenges than females ($M=1.77$, $SD=1.01$). There were no significant effects for the MGM group, but the results indicated a similar trend ($F(1,22)=3.94$, $p=0.06$) with males ($M=2.29$, $SD=0.73$) completing more challenges than females ($M=1.75$, $SD=0.62$)

The challenge scores were also examined separately for the two different goal groups. The data analysed was the best score that a dyad had attained across all challenges played. Choosing a dyad's worst or average score would be problematic because the pair may have run out of time at a point when they had attained an unrepresentatively low score for them. Again two 2x2 ANOVAs were performed to compare across SEX and COMM. None of the results were significant at the 0.05 level, though for the SGM condition there was a trend towards males ($M=81.85$, $SD=25.19$, $N=13$) scoring higher than females ($M=58.56$, $SD=39.35$, $N=13$) ($F(1,22)=2.88$, $p=0.10$).

8.1.3 *Relationship Between Game Performance and Academic Gain*

No. of chals (N)	*Best score*	*Improvement*	*Pre-test*
0 (3)	0.00	1.83	4.00
1 (4)	96.60	3.75	7.12
2 (17)	58.04	1.85	6.65
3 (20)	82.69	1.82	8.50
4 (6)	96.27	2.83	8.08

Table 8.4: Challenge score, improvement and pre-test as a function of number of challenges completed for SGM

Improvement and number of challenges completed

Tables 8.4 and 8.5 show how three other measures vary according to the number of challenges completed. The primary question addressed was whether academic improvement was a function of performance, with performance being measured in terms of number of challenges. Again, due to the fact that the two GOAL conditions were not comparable for performance measures, results were analysed separately for SGM and MGM players. As

No. of chals (N)	*Best score*	*Improvement*	*Pre-test*
0 (1)	0.00	-2.00	8.50
1 (6)	93.11	-1.83	9.50
2 (32)	88.67	0.10	9.50
3 (12)	82.30	1.67	8.71

Table 8.5: Challenge score, improvement and pre-test as a function of number of challenges completed for MGM

seen in the tables, results suggest that there is little relation between number of challenges completed and improvement for players in SGM, but that for MGM the improvement appears to increase along with number of challenges completed. Pearson Product moment correlations were performed to test these relationships, and it was found that there was no relationship for SGM (*r(48)=-0.12, ns, 2-tailed*), but that there was a significant positive correlation for MGM (*r(48)=0.31, p<0.05, 2-tailed*).

Challenge scores and pre-test totals
The two other scores given in Tables 8.4 and 8.5 are the best challenge score and the pre-test total. We were interested, first, in whether there was a relation between challenge scores and improvement, which correlations indicated was not the case for either GOAL mode. The second question regarding challenge scores was whether they were related to the other performance measure of number of challenges completed. As Table 8.4 indicates, a significant correlation was observed between challenge scores and number of challenges completed for SGM players (*r(24)=0.56, p<0.01, 2-tailed* [4]), indicating that those who completed more challenges also did well within at least one of the challenges they completed. For MGM players there was a similar positive but non-significant correlation. The fact that both of the correlations are positive discounts the possibility that doing more challenges may have led to a less optimal performance within challenge.

Regarding the pre-test data, we want to know if prior ability in the task-related area predicted performance in *Builder*. As suggested by the data presented in Tables 8.4 and 8.5, there was no correlation between pre-test total and number of challenges for either of the GOAL groups. There were, however, significant correlations between pre-test scores and challenge scores both for SGM (*r(49)=0.34, p<0.05*) and MGM players (*r(49)=0.36, p<0.01*).

Pre-test and improvement
An important complexity with the improvement scores was also discovered in that there was a strong negative correlation between pre-test and improvement scores (*r(98)=–0.62, p<0.01*). This may explain why achieving a good within-challenge score does not lead to a high improvement score. That is, those who went well did so partly because they had a good mastery of the domain, but their mastery prevented them from improving as much as

[4] The low number of observations is due to the fact that this correlation was conducted using the dyad as the unit of analysis, since both the measures are for game performance.

those with less mastery in the domain. Since the highest scores on the pre-test approached 100%, the lack of observed improvement in proficient students may also be explained by ceiling effects.

8.2 Sociomotivational Outcomes

Before looking at the questionnaire data, we analysed the results of the pre-game question, which asked players to rate on a 5-point scale how much they liked (playing with) their allocated partner. A SEX(M,F) x GOAL(SGM,MGM) x COMM(BCM,ECM) ANOVA revealed no significant main effects or interactions at the .05 level. This result alleviates the concern that some cells in the design may have been composed of closer friends than other cells, which could otherwise confound the results on attitude to partner.

The investigation of the questionnaire data began with a factor analysis on the 20 Likert-style attitude items contained therein. [5] The purpose of a factor analysis is to increase the robustness of the dependent measures by suggesting possible groupings (*factors*) of items according to their statistical relationships across observations. For the factors to be useful in analysis they must also be conceptually related. It is expected that this will be the case given that the items were designed to assess a particular set of constructs.

The Principal Components Analysis produced six factors with eigenvalues over 1.0, a typical selection criteria for factors worthy of further analysis. Three of these factors were excluded on the basis of investigation of the Scree Plot, which tends towards a straight line after the third factor, suggesting that only the top three factors accounted for sufficient variance to be worth considering. Factors 1, 2 and 3, respectively, accounted for 27.1%, 11.8% and 9.4% of the variance, making a total of 48.3% accounted for. Below are presented the factor loadings of each item for the three factors, as well as the mean item scores.

Recalling that a "yes" is 4.0 on the scale, and a "YES" is 5.0, inspection of the means of Table 8.6 indicates players reported generally positive attitudes. The highest rated items were: Q1 ("I enjoyed playing *Builder*", $\underline{M}=4.55$), Q8 ("I would like to play *Builder* again", $\underline{M}=4.55$) and Q17 ("I wish I could have played *Builder* for longer", $\underline{M}=4.43$). Although positive attitudes are good, the high means and relatively low standard deviations raise the concern that scores will not discriminate between conditions well because there is not a large enough range of results. It could be hypothesized that the high scores were partly a result of respondents trying to please researchers.

Inspection of the loadings onto Factor 1 indicates that over half of the items are above the typical criteria of 0.4, possibly due to the uniformness of answers discussed above. It would be difficult to conceptually interpret such a large group of items. Factors 2 and 3, on the other hand, provide sets of appropriately related items. The items loading onto Factor 2 were: Q1 ("I enjoyed playing *Builder*"), Q10 ("I would like to play *Builder* at home"), Q17 ("I wish I could have played *Builder* for longer"), Q4 ("Computer games like *Builder* should be used more in school"), and Q8 ("I would like to play *Builder* again"). These can be sensibly grouped under the heading *persistence of interest in activity*, one of the target areas of the questionnaire. The items loading onto Factor 3 were: Q12 ("If I play *Builder* again I

[5] The final three questions of the questionnaire were omitted as they did not relate to the game, but to the respondent's everyday patterns of computer use.

	Factor 1	Factor 2	Factor 3	Mean	SD
QUEST1 [enjoy game]	.67525	.47009	.00525	4.55	0.57
QUEST10 [home game]	.58433	.58240	-.02327	4.15	0.98
QUEST12 [again partn]	.65602	-.25319	.44500	3.69	1.11
QUEST13 [other partn]	.42513	-.09363	.48746	3.78	1.16
QUEST14 [like comm]	.61700	-.22693	-.27850	3.92	0.81
QUEST16 [same room]	.37022	-.20292	-.46613	3.69	1.35
QUEST17 [play longer]	.45604	.45428	-.28802	4.43	0.81
QUEST18 [easy comm]	.58103	-.37145	-.52007	4.02	0.89
QUEST2 [like comp]	.26929	.38001	.14002	4.39	0.62
QUEST19 [prefer alone]	.55895	-.24421	-.03429	4.29	0.93
QUEST20 [comm helped]	.63982	-.27900	.13648	4.46	0.84
QUEST3 [friend partn]	.40344	-.16561	.47171	4.36	1.00
QUEST4 [school game]	.39073	.50203	.04354	4.20	0.97
QUEST5 [diffic comm]	.45664	-.34204	-.44847	3.78	1.20
QUEST7 [partn game]	.64526	-.13592	.14240	4.43	0.81
QUEST8 [again game]	.64868	.61452	-.16597	4.55	0.71
QUEST9 [help partn]	.49936	-.32521	.43879	4.02	1.09
QUEST6 [easy game]	.15958	-.00837	-.02034	3.54	0.87
QUEST15 [more instrn]	.53028	-.32543	-.24217	3.92	1.13
QUEST11 [learn game]	.54092	.06815	.20106	3.81	0.95

Table 8.6: Factor loadings and means of questionnaire items

would like to play with the same partner"), Q13 ("I would enjoy playing other games with the same partner"), Q3 ("My partner was friendly") and Q9 ("My partner was helpful"). Therefore this factor can be considered as measuring *affect towards partner*. Given that factors 2 and 3 address two of the four areas that the questionnaire aimed to assess, Factor 1 was considered in relation to the other two areas: *attitude towards the game* and *perception of collaboration*. Q14 ("I liked communicating with my partner using the computer"), Q18 ("It was easy communicating with my partner using the computer"), Q19 ("I would prefer to play *Builder* alone" [reverse scored]), Q20 ("Communicating with my partner helped us play the game") and Q7 ("I like playing computer games with a partner"), were all amongst the highest of the loadings on Factor 1. According to purely statistical criteria, the other items that loaded highly onto Factor 1 should be included. However, so that the group of items was interpretable, all those items that cross-loaded onto Factor 2 or Factor 3 were excluded, as were those which did not relate to attitude towards communication, based on conceptual grounds. Given these mixed criteria for inclusion of Factor 1 items, it was decided to consider separately one item relating to the perception of collaboration. At the time of developing the questionnaire it was felt that responses to Q16 ("I would prefer to have my partner in the same room") might be particularly interesting, partly because the item addresses the central CSCL/HCI question of simulating real collaboration in a virtual setting, but also because the wording makes the item less open to the problem of respondents trying to please researchers (at least compared to items like Q14, "I liked communicating using the computer"). The fact that the mean was relatively low for the question also suggested that it may be somewhat more useful as a distinguisher. Therefore Q16 was included separately in the analysis.

For the items returned for Factors 1, 2 and 3, a total score was computed as the average score received across items included in each factor. Higher scores always indicate a more positive attitude. A series of four SEX(M,F) x GOAL(SGM,MGM) x COMM(BCM,ECM) ANOVAs were performed for Q16 and each of the three factors. Prior to this, the relationship between partners' scores on each of the measures was considered to decide whether the analysis should be conducted at the individual or dyad level.

Perception of collaboration (Factor 1)

Results of the correlational analyses revealed a significant relationship ($r(48)=0.34$, $p<0.05$) between partners on Factor 1, so the dyad was used as the unit of analysis. There were no significant main effects in the ANOVA, but there was an interaction between SEX and COMM ($F(1,42)=4.70$, $p<0.05$). As suggested by the means of Table 8.7, post-hoc analyses (Tukey) revealed that females were significantly more positive when in ECM than in BCM ($F(1,49)=4.869$, $p<0.05$), whereas the two conditions did not differ significantly for males.

Persistence of interest in game (Factor 2)

Factor 2 scores did not show a significant correlation between partners ($r(48)=0.22$, *ns*), so the unit of analysis was the individual player. Results of the ANOVA revealed a significant main effect for SEX ($F(1,91)=4.00$, $p<0.05$), with males ($\underline{M}=4.43$, $SD=0.54$) scoring higher than females($\underline{M}=4.17$, $SD=0.87$), and a significant GOAL*SEX interaction ($F(1,91)=5.22$, $p<0.05$). Post-hoc analyses (Tukey) revealed that males in SGM showed significantly greater persistence of interest than females in SGM ($F(1,46)=7.43$, $p<0.01$), but there was no parallel gender difference in MGM (see Table 8.8).

		BCM	ECM
Male	*M*	4.33	4.15
	SD	0.57	0.46
	N (pairs)	13	12
Female	*M*	4.08	4.5
	SD	0.62	0.39
	N (pairs)	12	13

Table 8.7: Means for Factor 1 (Perception of Collaboration) showing SEX*COMM interaction

		SGM	MGM
Male	*M*	4.64	4.27
	SD	0.30	0.63
	N (pairs)	23	26
Female	*M*	4.05	4.31
	SD	0.95	0.66
	N (pairs)	26	24

Table 8.8: Means for Factor 2 showing SEX*GOAL interaction

Attitude to partner (Factor 3)
The correlation between partners' scores on Factor 3 indicated that the dyad should be the unit of analysis ($r(48)=0.41$, $p<0.01$). Results of the ANOVA again revealed a significant main effect for SEX ($F(1,42)=4.78$, $p<0.05$), with females ($M=4.05$, $SD=0.46$, $N=25$) being more positive towards their partners than males ($M=3.75$, $SD=0.97$, $N=25$). There was also a significant GOAL*COMM interaction ($F(1,42)=5.78$, $p<0.05$) for which none of the post-hoc tests were significant at the 0.05 level. A notable trend which approached significance was the difference between the two GOAL modes with COMM held constant at BCM ($F(1,23)=3.46$, $p=0.07$), indicating that given basic means of communication, MGM players were less positive about their partners than SGM players (see Table 8.9).
Same room (Question 16)
Individual scores were the unit of analysis for Q16 as there was no significant correlation between partners ($r(48)=0.08$, *ns*). The ANOVA revealed a significant main effect for COMM ($F(1,92)=4.19$, $p<0.05$), with those in ECM ($M=3.96$, $SD=1.19$, $N=50$) scoring higher than those in BCM ($M=3.38$, $SD=1.48$, $N=50$). The intent of this question was to ascertain how frustrated players became by not being able to communicate face-to-face. It was reverse-scored because a high desire to have the partner in the same room indicates a negative attitude towards the communication. Hence the results indicate that those in

		SGM	MGM
BCM	*M*	4.18	3.63
	SD	0.46	0.90
	N (pairs)	12	13
ECM	*M*	3.85	4.14
	SD	0.58	0.60
	N (pairs)	12	13

Table 8.9: Means for Factor 3 showing GOAL*COMM interaction

ECM felt less frustrated in their attempts to communicate than those in BCM.

8.3 Game-play

One of the main points of interest in relation to game-play was the communication between players. This section presents data on the volume of communication between players of different conditions, as well as some anecdotal examples of interesting features of communication and other aspects of game-play.

Number of written messages sent	
Males: 27.34	*Females:* 28.44
SGM: 28.45	*MGM:* 27.31
BCM: 35.6	*ECM:* 20.02

Table 8.10: Written message count across SEX, GOAL and COMM

Amount of sound data sent (in Kilobytes)	
Males: 175.5	*Females:* 183.6
SGM: 157.6	*MGM:* 200.8

Table 8.11: Volume of spoken communication across SEX and GOAL

Tables 8.10 and 8.11 present data from the log files regarding number of written messages and volume of spoken data sent between partners during game-play. Table 8.10 includes the means for the two COMM groups, showing that BCM players wrote more messages to each other than ECM players. Given that there is less need to write messages if players have speech, this difference is to be expected, and COMM was not included in the ANOVAs. A second expectation, given the findings in the computer game literature that females are more interested in games that involve relationships and communication, was that females would communicate more than males, which does not appear to be supported by the means. The other hypothesis with regard to communication was that a less-structured task (MGM) might lead to more communication. The means for speech data show some support for this hypothesis, with MGM players sending an average of 200KB per session compared to SGM's 158KB. SEX(M,F) x GOAL(SGM,MGM) ANOVAs, however, revealed no significant main effects or interactions in either the written or the spoken data. One of the more interesting points of the communication data analysis is that males and females differed so little on these measures. This was particularly surprising in the light of observations that suggested females were more interested in the communication than males. A tentative interpretation of the lack of difference is that the challenge of the game itself led to males communicating more, even though they did not appear to be as interested in the communication as in completing the challenges.

Observation Forms

Data from the observation forms is important because it provides information not obtainable from the written tests, questionnaires or the log files. For example, there was no indication in the log files of what players in ECM talked about. There was also no way to ascertain how comfortable players were with the interface. Although data from the observation forms is subjective and variable in level of detail, it is helpful when used in light of the more formal results – particularly if observations can be used to suggest explanations of the other results. The observations recorded included the following types of information:

- whether players had trouble with the interface;
- how well players worked together;
- how focussed players were on the task; and
- whether players had a good sense of the concept.

Several players were observed to be either particularly adept or inept with respect to the interface of *Builder*. There was no noticeable tendency, however, for any particular group to fall into either of these categories. In particular, there was no evidence from the observations that girls were any worse at adapting to the interface than boys, which might be expected based on the reported gender differences in computer proficiency discussed in the literature review.

Concerning the observed degree of collaboration, more male pairs were noted to have collaborated ineffectively than female pairs, while there was no difference in which pairs were noted to have collaborated well. Comments from the observation forms regarding male pairs collaborating poorly included: "didn't seem to be working together", "poor communication with partner", etc.

More male pairs were noted to have shown a strong grasp of the concept or to be particularly focussed on achieving the goal. It was also noted that more players in MGM had trouble with the concepts than players in SGM.

8.4 Reported Home Computer Use

The final three items on the questionnaire were open-ended questions regarding the participants' home computer use. There were two motivations for asking these questions. The first was to gather descriptive data about computer use amongst school children of the target age group, and the second was to compare use across gender. The other, manipulated independent variables of the design are not relevant for these questions, since home computer use is outside of the context of the study. Results for the three questions were as follows.

- **Number of computers at home:** The overall mean for this item was 1.65 computers (*N=99, SD=0.93*). Only three respondents reported having no computers at home. [6] The most common responses were one computer (49 respondents) and two computers (30). A slightly higher number of computers was reported by females (*M=1.72, SD=0.94, N=48*) compared to males (*M=1.59, SD=0.92, N=51*), though this was not significant according to a *t-test* (*t(97)=-0.70, ns*).

- **Hours per day spent on computer (at home):** The average amount of time spent using a computer per day was 1.13 hours (*N=101, SD=0.92*). The most common responses were one hour (30), half an hour (15), two hours (14) and 1.5 hours (12). There was a non-significant trend for males (*M=1.28, SD=1.07, N=52*) to spend longer on the computer than females (*M=0.97, SD=0.70, N=49*) (*t(99)=1.69, p=0.09*).

- **What respondents used the computer for:** This item produced an incredible range of responses, including over 50 different games. Other than games, respondents reported using computers for typing, painting, homework, writing letters and email, using the internet, and using some other form of reference material, mostly electronic encyclopedias. Reports of "typing", where the exact purpose of the typing was not specified, were counted in the same category as doing school homework or projects. Table 8.12 indicates the number of males and female respondents who reported computer use within each of the usage categories. The results for non-game uses are presented first. To summarize the game responses we decided to present results only for the games mentioned by more than two females or more than two males.

  Aside from the games presented in the table, a casual glance at the names of other games listed by males and females supports the general distinction. Most of the games listed by male respondents were either sports-related (e.g., Lakers vs. Celtics, Micheal Jordan In Flight, NBA) or were among the dizzying array of fantasy/battle-oriented games available on the market (e.g., Time Commando, Dune 2, Star Wars, Wing Commander, Crusader, Alpha Battle, Mech Warrior, Quake, Hellbender, etc.). The games listed by girls tended to be adventure-oriented or based on popular television or movie themes (e.g., Carmen Sandiego, Alladin, Lion King, Pumba and Timon's Jungle Games, Treasure Cove, Simpsons, Dino Park, Dick Tracy, Theme Park). There was a good deal of gender overlap on certain types of games, most notably card games and SimCity, but the expected imbalance with regard to violent action games is clearly

[6] It should be noted that, in addition to the three respondents who reported zero computers, there were three missing values, which may have been intended to indicate zero.

	Females	Males
Homework	26	28
Letters	2	0
Stories	2	1
Email	4	1
Encyclopedia	5	6
Internet (general)	17	13
Internet (chat)	3	1
Painting	6	2
Card games	12	6
Tetris	6	1
Pinball	5	3
JezzBall	4	3
Yukon/Amazon Trail	3	0
CC Red Alert	1	15
SimCity	4	8
Doom	1	7
NHL 97	1	7
Warcraft	0	6
Duke-Nukem	1	6
Car racing	2	5
Golf	1	3

Table 8.12: Uses of computer across gender

demonstrated, particularly with respect to Command and Conquer Red Alert, Doom and Duke-Nukem. [7]

[7]Referring back to the discussion in the literature review, Duke-Nukem is a prime example of a male-biased game, particularly in terms of its incidental objectification of women.

Chapter 9

Discussion

9.1 Summary of Findings

9.1.1 *Achievement*

To begin with the academic outcomes, the results revealed two significant effects. First, in the control-group comparison, it was found that pairs playing the remote CSCL activity *Builder* showed greater improvement in the target mathematical areas than students who received no instruction. Second, it was found that the manipulation of the goal specification within the activity affected the degree of improvement. Specifically, players who were given a clear, numeric area or volume goal showed greater academic improvement than players who were given the more open-ended goal of maximizing the area or volume. Results for the second manipulated variable, mode of communication, failed to show that enhancing communication with spoken messages and virtual presence led to greater achievement. There were no significant gender differences on academic improvement, but it was found that males completed more challenges and scored higher in the activity than females. There was a positive relationship between number of challenges completed and academic improvement, but only for players in *maximize goal mode* (MGM).

9.1.2 *Attitudes*

There were several communication- and gender-related differences found in the sociomotivational outcomes. Males showed a greater persistence of interest in the game than females, suggesting that the activity was more motivating for males. This gender difference in attitude to the game, however, was not uniform across modes of task, as the analysis of the interaction with the task mode revealed that males showed significantly greater persistence of interest than females only when in *specific goal mode* (SGM). In terms of perceived collaboration, results indicated that females were more positive about the communication when they had speech and virtual presence (ECM), while for males this did not effect their attitude. Furthermore, responses to the specific item "I would prefer to have my partner in the same room" showed the expected result that players with *enhanced communication* (ECM) found the remote collaboration less frustrating than players with *basic communication* (BCM). Finally, in terms of the social measure of interpersonal attitude, it was found that females were more positive about their partners than males. Given that there was no parallel gender difference on the pre-game rating of attitude toward partner, the results imply that the game had a positive effect on interpersonal attitudes (as expected from the Cooperative Learning literature), but only amongst females. In terms of the manipulated variables, SGM generated a more positive attitude to the partner than MGM when players were in *basic communication mode* (BCM).

Anecdotal observations supported the performance findings in that more males were observed to be strongly focussed on the task than females. There was also anecdotal evidence of some females being more interested in communicating than in the task, compared to males, and there were more observations of males communicating ineffectively than females. Players in MGM were observed to have more trouble with the game than players in SGM, which may help explain the achievement findings.

9.2 Relating Findings to the Literature

9.2.1 *Remote Collaborative Learning?*

When introducing the motivations for the study, two categories of interest were identified, the first of which was the question of whether the positive results of Cooperative Learning can be facilitated in a remote computer-supported setting. As previously stated, the design of the current study did not attempt to answer this question in terms of whether remote CSCL is better than individual CAL, co-present CSCL, or non-computer Cooperative Learning. While these questions could foster interesting follow-up research, the position advocated in the current study is that such comparisons may be ungrounded in terms of the educational goals of the different methods, and may also be difficult to investigate experimentally. Nevertheless, on the basis of the control group comparison and some of the descriptive data presented in the results, several observations can be made regarding this broad question.

First, there is evidence that the *Builder* activity led to significant learning gains in the target mathematical areas. Why did *Builder* produce this improvement? One explanation is that players were able to rehearse existing knowledge by applying it to the problems presented in the activity. Given the collaborative nature of the game, it may also be that players improved because they had to verbalize their problem-solving strategies. Such verbalization, and consequent rehearsal and reappraisal of learning constructs, has been found to be the primary agent of academic improvement in CL. Based on the current evidence, it is not clear whether the task itself or the collaboration played the greater role in improvement. Looked at from the learners' perspective, however, the collaboration appeared to play a central role. This was evident in that the third highest-ranked item in the questionnaire (with a mean of 4.46) was "Communicating with my partner helped us play the game".

Second, the results of the attitudinal questionnaire suggest that *Builder* has a positive effect in non-academic areas. Those who played the game appeared to find it substantially motivating and engaging; as indicated by the two highest-ranked items on the questionnaire, which were "I enjoyed playing Builder" and "I would like to play Builder again" (each with a mean score of 4.55 on the 5-point scale). The positive mood evoked by the game may have transferred to the isometric activity of the post-test, providing an alternative explanation for the improvement seen in the game-group. The magnitude of this general positive attitude toward the game is difficult to ascertain, since we are not comparing those who played the game with those who performed some other activity. There is also the possibility that other aspects of the study could have contributed to the positive responses; one example being that students may have enjoyed the variation from normal classwork, and that given the short length of play there was not long enough for the initial novelty of the game to wear off. Further studies within slightly different contexts, such as having students play *Builder* in their lunch break everyday for a week, would help confirm the degree of generated engagement.

Nevertheless, the positive results discovered were what we expected, given that the activity was designed to be game-like (i.e., more like play than work), and they were supported by anecdotal observations, both in the final study and the pilots. It was noted that communication was a substantial component in students' appreciation for the game, with those who had only BCM often expressing disappointment when they later found out that other players could talk to each other. While it was found that almost all players enjoyed the communication, there were more noticeable individual differences in how absorbed different players became in the activity itself. Several players – especially, but not only, boys – worked almost feverishly on the challenges during their session, and expressed a desire to play *Builder* again during break time. There were a few players, however, who seemed either confused by or relatively uninterested in the activity – this was more often the case for players in MGM. In terms of the varying attitudes toward *Builder*, the results of the current study are illuminating mainly in regard to gender, as is discussed further below. These additional anecdotal observations indicate that it may be worthwhile to further investigate different types of learners to decide who "profits" most (both in terms of engagement and achievement) from a specific CSCL tool. A non-trivial challenge that must be addressed by researchers in the field is that games and collaboration should involve and stimulate interest in learners who are not already proficient in or motivated by the target learning area. Such research requires extensive profiling of the learner, beyond simple distinctions such as gender or academic level.

In summary, the academic control-group comparison and descriptive data on general attitude to the task indicate that *Builder* may be quite effective in the role for educational games proposed by [KP95] – that of stimulating interest within and supporting a broader instructional environment. The findings are also an encouraging step towards incorporating computer-supported collaboration into Distance Education. Future study should assess what factors mediate engagement across different types of learner, so that alternative interface or task elements can be considered to include a wider range of learners.

9.2.2 *The Role of the Task*

The most important variable manipulated within the activity in terms of academic improvement was the nature of the task. Given the apparent conflict between the current results and the success of ill-structured tasks in Cooperative Learning [Coh94, JMJ+81], we need to consider why the clearly-structured goal had more effect on learning. An initial consideration in response to this conflict is that the SGM task, though having a well-defined goal, was still ill-structured to the extent that there are many possible solutions to a challenge and many decisions that players may wish to discuss. The remainder of this section suggests other explanations for the success of SGM.

Differences between the two task modes that may be responsible for the results include: the fact that SGM has a clearer goal; the direct feedback in SGM; and the relative difficulty of performing optimally in MGM. Having a clear goal has been found to be the most important deciding factor in the popularity of games [Mal82], hence it may have been more engaging for learners. Immediate and simple computer feedback has also been found to be more motivating than less-tangible, self-regulated feedback in co-present CSCL [NC93].

More specific features of the current context may also explain the results. If MGM is a harder activity, some players may have failed to grasp the concepts, or at least had trouble doing so during the short playing time. There was support for this in anecdotal observations made by the researchers that more players in MGM appeared to be struggling or confused by the game. From observations of others who have played the game outside of the formal study, it would appear that for older players MGM is more challenging and interesting than SGM. The direct feedback of SGM may also be less important for adults ([NC93]'s study was also with elementary-level students). The relevance of age is supported by comparison with [BTR+95]'s multi-input CSCL study of high-school and college students using a simple activity with a clear goal. Collaboration had no significant learning effect, and many players in the college student sample said that they would prefer to play alone. The type of activity used by [BTR+95] may work better with a population of younger learners, while tasks such as MGM in Builder may work better with older learners. Future research may benefit from addressing such developmental issues.

The task results may also be domain-dependent. In SGM there is a greater need for players to be focussed on the numbers, and this emphasis on detail may be particularly beneficial in mathematical learning. There is a need for further research to ascertain the effect of ill-structured goals in other domains (such as the languages and social sciences). To conclude, the results obtained indicate that within a short time period, the use of a CSCL activity with a specific goal can have a postitive effect on task-related learning in mathematics for young learners.

9.2.3 *Supporting Interaction*

Of equal interest to the significant effect found for the goal manipulation itself is the fact that it was more influential than the mode of communication. It was expected, based on the elemental role of learner interaction in Cooperative Learning, that varying communication would have a greater impact than varying the specificity of the task. The straightforward interpretation of the lack of effect is that learners found they could communicate as well with writing as they could with speaking. To further illuminate the role of communication in the remote CSCL setting, a follow-up study comparing spoken and written communication with no verbal communication is necessary. In the latter condition, players would be forced to rely purely on aspects of the visual interface in order to collaborate.

An alternative explanation for the equivalence of ECM and BCM in the academic results is that students did not feel comfortable using the speech because it was embarrassing. This reason was volunteered by some students, either spontaneously or when asked why they used speech sparingly. Furthermore, it may be that learning improvements emerged more out of the task itself than the communication between players, which could be addressed by comparing the current outcomes with those from a single-player version of the game.

Although the different modes of communication did not affect academic improvement, they had significant effects on the sociomotivational outcomes, a result also expected from the CL literature. Responses to the "same room" question indicated that ECM made it significantly easier for players to collaborate remotely. Females in ECM were also more positive about the collaboration in general than those in BCM. Enhancements to the communication were therefore important in terms of perceived collaboration.

The interaction of the effects of task and communication in the attitude data is suggestive of the role different modes of communication play in different types of CSCL activities. It was found that players in MGM were less positive about their partners if they had only basic communication, suggesting there may be more of a need for enhanced communication within less-structured activities. Within a certain range of domains (e.g. mathematics) and learning activities (e.g. well-structured), written communication may be as effective as multiple channels of communication. On one hand we can see these results as positive indications for distance education projects that use simple forms of communication, such as email, bulletin boards or chat facilities. These styles might be quite sufficient for collaborative learning within certain domains and types of task. On the other hand, the results suggest the need for further work in enhancing communication to support less-structured tasks. Furthermore, the attitudinal results suggest that to create a subjective environment of collaboration and positive regard for co-learners, enhancing communication with spoken communication and/or virtual presence may be benificial.

9.2.4 *Gender Differences*

While there were no significant gender differences in academic improvement, it was found that males completed more challenges in *Builder*, and had different attitudes towards the game, their partners, and collaboration in the game. The fact that males showed a greater persistence of interest in the game than females, particularly when in SGM, supports findings from previous studies indicating that goal-focussed games with record scores are more interesting to boys than girls [OKdV96]. The alternative hypothesis, that the prominent role of communication might make the game more interesting to girls, was not supported by the results. Other attitude measures discussed above also showed gender differences. First, females were more positive about the communication when they had speech and virtual presence, while for males this did not affect their attitude. Second, in terms of the social measure of interpersonal affect, it was found that the game had a more positive influence on females than males.

A general way to summarize these findings would be to say that males responded well to having a specific goal, while females responded well to being able to speak to and/or see an image of their partner. This was supported by the anecdotal observations of males being more strongly focussed on the task and females being more interested in communicating. Findings such as these gender differences can be used as guidelines in the design of learning activities like *Builder* to ensure the inclusion of elements that are effective for different types of learners.

9.3 Future Research

The current results suggest guidelines for future research to further elucidate the important elements in a remote CSCL setting. One suggestion is to investigate whether the positive effect of a specific goal and the lack of effect of enhanced communication is replicable in different domains, with learners of different ages, and over longer periods of instruction. Two specific questions that have emerged in relation to learning are:

- Is there a set of learners, a domain, and/or a duration over which less-structured goals show greater academic gain in remote CSCL?
- Does enhancing communication within the less-structured setting influence the size of the gain?

With respect to the gender results, the most important consideration is whether there is a real benefit in different styles of communication and task for different types of learners. One interesting approach to this would be to create an entirely user-configurable learning environment, such that users can choose both how structured their goal is and what type of communication channels they use. Such an open-ended investigation would help define the set of elements within remote CSCL that are valuable across the range of possible learners.

Interpreting the results in the context of related work suggests a wide range of directions that could be investigated in further studies with the *Builder* activity. For example, we could study *Builder* when played within the broader, MUD-like setting of *Island*. This would involve players making more decisions about how to spend their time, which might require more or different types of interaction. Within a large game of *Island* the notion of group decisions or politics could be introduced. The more decisions that have to be made, the more opportunity there is for divergent ideas and conflict, factors claimed to be powerful agents in cognitive change by Nastasi and Clements. Modes of interaction could be manipulated within such a setting, as a supplement to the study of the lower-level question regarding media of communication. Elements of the interaction could also be used as dependent measures, and again the observational assessments used by Nastasi and Clements [NC93] could provide important information supplemental to the learners' questionnaire responses.

Placing *Builder* within *Island* also provides a different type of motivation to perform well in the activity. A player's house would be displayed to all the other players in the broader game, and might also have to withstand weather or other hazards. It may be that some learners are more motivated by such concerns. This enriching of the fantasy aspect of the game is in accordance with the work of Malone.

Further studies could alternatively concentrate on the aspects of the interface that support learner interaction. The investigation of virtual presence, for example, could be extended to allow a limited set of symbolic gestures, which could be compared with the provision of a more open-ended tool for gesturing (such as manipulating simulated 3-D hands). How users choose to represent themselves in a fantasy-type virtual environment is also an interesting research question, and one that is timely in the context of the emerging cyberspace technology. In the current study some gender differences were observed in type of icon chosen, but this could be further explored by allowing a much broader range of choices, the adoption of names, or even allowing players to create their own avatars. The anonymity and possibility for identity creation may lead to interesting types of collaboration not found within co-present models.

Appendix A

Implementation Details

CodeWarrior (versions 9 and 10) was used as the compiler and debugger. ResEdit was used for handling resources (pictures, sounds, etc.). The free-ware library SpriteWorld was used for many of the graphical objects. This was originally written by Tony Myles, but has been updated by Karl Bunker and Vern Jensen (official internet site: `http://members.aol.com/SpriteWld2/index.html`). Brett Allen, an E-GEMS programmer, wrote an add-on to this library called SpriteCan, which was also used in *Builder*. Thanks are also due to Brett for help with trouble-shooting, particularly with SpriteWorld-related issues. Other credits are due to the following E-GEMS programmers. Greg Smolyn wrote the `Client` and `Server` classes for the overall game *Island* (this code was adapted from the previous *Island* driver written by Hardeep Sidhu and myself). Greg also wrote the 3D renderer used in *Builder*. Omar Odeh wrote the original network code and also the functions dealing with the written messages. Ben Walton wrote the code for handling the log files and the high score files. Ryan Fugger provided the sound effects for Builder and also wrote the functions dealing with spoken messages. Graphics for the start-up screens for Builder were provided by Fiona Wong.

Note that the communication and task modes (i.e., whether players are in ECM or BCM, and SGM or MGM respectively) are set at compile-time by flags in the source code, meaning they cannot be changed during run-time.

Appendix B

Client-Server Messages

Messages are sent between the `BuilderServer` and the `BuilderClient` using the generic *Island* message package, which includes a specific sub-package called "BuildNetMessage". The header file BuildNet.h declares a BuildNetMessage as follows:

```
struct BuildNetMessage{
        netEvtType what;
        short field1;
        short field2;
        short field3;
        Point where;
        Rect inf;
        char special[1];
};
```

The message type, which is the first thing looked at by `BuilderServer` or `BuilderClient` upon receipt of a message, is given in the "what" field, which is of type "netEvtType". The `enum` statement defining netEvtType is:

```
enum netEvtType{ kChooseRequ, kNewWRequ, kLockRequ, kChangeWallRequ,
                 kChooseConf, kNewWConf, kLockConf, kChangeWallConf,
                 kRendInRequ, kAreaRequ, kChangeSideRequ, kEndRequ,
                 kRendInConf, kAreaConf, kChangeSideConf, kEndConf,
                 kVPRequ, kPerspRequ, kJoinRequ, kRendMoveRequ,
 kVPConf, kPerspConf, kJoinConf, kRendMoveConf,
 kDoneJoin, kBrickConf, kFrameInit, kFrameTake, kFrameReplace };
```

As will be noted, most message types end in either "Requ" or "Conf". The prior indicates a request from the client, while the latter is a confirmation from the server (which is often sent to both clients). The meaning of each type is briefly stated below. In the cases where there is both a request and a confirmation of the same type, only the request is defined.

- `kChooseRequ` – request to set the challenge (from "Challenge-Selection" screen).

- `kNewWRequ` – request to make a new wall.
- `kLockRequ` – request to select a wall.
- `kChangeWallRequ` – request to move, flip, resize or delete a wall.
- `kRendInRequ` – notification of entry to 3D-View (other client has to be informed).
- `kAreaRequ` – request for current area/volume calculation.
- `kChangeSideRequ` – request to move, delete or set frame of a window or door.
- `kEndRequ` – request to quit a challenge.
- `kVPRequ` – request to set icon (from "Challenge-Selection" screen).
- `kPerspRequ` – notification of switch between Top- and Side-View.
- `kJoinRequ` – new client's request to join *Builder*.
- `kRendMoveRequ` – notification of movement in 3D-View.
- `kDoneJoin` – server tells new client that initializing information is complete.
- `kBrickConf` – server updates client on current brick count.
- `kFrameInit` – server gives client list of frames allocated at the start of a challenge.
- `kFrameTake` – server informs client of a frame removed from the store.
- `kFrameReplace` – server informs client of a frame returned to the store.

Appendix C

Academic Tests

C.1 Pre-test

.. attached ..

C.2 Post-test

.. attached ..

Appendix D

Questionnaire

For each item 1–20 players were given a set of 5 words to choose from: NO, No, Maybe, Yes, YES.

1. I enjoyed playing Builder.
2. I like using computers.
3. My partner was friendly.
4. Computer games like Builder should be used more in school.
5. It was difficult to communicate with my partner.
6. I thought the game was easy to play.
7. I like playing computer games with a partner.
8. I would like to play Builder again.
9. My partner was helpful.
10. I would like to play Builder at home.
11. I learned something by playing Builder.
12. If I play Builder again I would like to play with the same partner.
13. I would enjoy playing other games with the same partner.
14. I liked communicating with my partner using the computer.
15. I needed more instructions to play Builder.
16. I would prefer to have my partner in the same room.
17. I wish I could have played Builder for longer.
18. It was easy communicating with my partner using the computer.

19. I would prefer to play Builder alone.
20. Communicating with my partner helped us to play the game.
21. How many computers do you have at your house ?
22. About how many hours a day do you think you spend on a computer (don't include school) ?
23. Please make a list of the main things you do on computer. If you play games, please write the names of the games you play.

Appendix E

Observation Form

.. attached ..

Appendix F

Builder Screenshots

.. attached ..

Bibliography

[ARS97] John R. Anderson, Lynne M. Reder, and Herbert A. Simon. Applications and misapplications of cognitive psychology to mathematics education. *(in press)*, 1997.

[Bat92] Trent Batson. Finding value in CSCL. *SIGCUE Outlook*, 21(3):26–28, 1992.

[Bro93] Herb Brody. Video games that teach? *Technology Review*, November/ December:51–57, 1993.

[BTR+95] Lauren J. Bricker, Steve L. Tanimoto, A. I. Rothenberg, D. C. Hutama, and T. H. Wong. Multiuplayer activities that develop mathematical coordination. In *CSCL '95: First International Conference on Computer Support for Collaborative Learning*, pages 32–39, 1995.

[Bul94] W.M. Bulkeley. Gender affects how user sees the computer. *Wall Street Journal*, March 16:B1–B4, 1994.

[Car87] John M. Carroll. The adventure of getting to know a computer. In William S. Buxton Ronald M. Baecker, editor, *Readings in Human-Computer Interaction: A Multidiscliplinary Approach*, pages 639–648. Morgan Kaufmann, San Mateo, CA, 1987.

[Cle81] Frank J. Clement. Affective considerations in computer-based education. *Educational Technology*, April:28–32, 1981.

[Coh94] Elizabeth G. Cohen. Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1):1–35, 1994.

[CToBS88] Multilevel Edition Canadian Tests of Basic Skills. Nelson Canada, Englewood Cliffs, NJ, 1988.

[Dal90] David W. Dalton. The effects of cooperative learning strategies on achievement and attitudes during interactive video. *Journal of Computer-Based Instruction*, 17(1):8–16, 1990.

[DG95] D.Griffin and R. Gonzalez. Correlational analysis of dyad-level data in the exchangable case. *Psychological Bulletin*, 118(3):430–439, 1995.

[DHH89] David W. Dalton, Michael J. Hannafin, and Simon Hooper. Effects of individual and cooperative computer-assisted instruction on student performance and attitudes. *Educational Technology Research and Development*, 37(2):15–24, 1989.

[Gol92] Shelley V. Goldman. Computer resources for supporting student conversations about science concepts. *SIGCUE Outlook*, 21(3):4–7, 1992.

[Gol96] Murray W. Goldberg. CALOS: An experiment with computer-aided learning for operating systems. In *ACM's 27th SIGCSE Technical Symposium on Computer Science Education*, pages 175–179, Philadelphia, PA, 1996.

[GRWB92] S. Greenberg, M. Roseman, D. Webster, and R. Bohnet. Human and technical factors of distributed group drawing tools. *Interacting with Computers*, 4(3):364–392, 1992.

[HDH87] Micheal H. Hannafin, David W. Dalton, and Simon Hooper. Computers in education: Ten myths and ten needs. *Educational Technology*, October:8–14, 1987.

[HH91] Simon Hooper and Michael J. Hannafin. The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development*, 39(3):27–40, 1991.

[HLKM92] R. Hertz-Lazarowitz, V. B. Kirkus, and N. Miller. Implications of current research on cooperative interaction for classroom application. In R. Hertz-Lazarowitz and N. Miller, editors, *Interaction in cooperative groups: the theoretical anatomy of group learning*. Cambridge University Press, New York, NY, 1992.

[HSGB82] Jan Hawkins, Karen Sheingold, Maryl Gearhart, and Chana Berger. Microcomputers in schools: Impact on the social life of elementary classrooms. *Journal of Applied Developmental Psychology*, 3(4):361–373, 1982.

[HZD93] Shelley Hymel, Beverly Zinc, and Elise Ditner. Cooperation versus competition in the classroom. *Exceptionality Education Canada*, 3(1-2):103–128, 1993.

[IBGK95] Kori M. Inkpen, Kellogg S. Booth, Steven D. Gribble, and Maria M. Klawe. Give and take: Children collaborating on one computer. In J.M. Bowers and S.D. Benford, editors, *CHI '95: Human Factors in Computing Systems*, pages 258–259, Denver, CO, 1995. ACM. Conference Companion.

[IBK95] Kori M. Inkpen, Kellogg S. Booth, and Maria M. Klawe. Playing together beats playing apart, especially for girls. In *CSCL '95: First International Conference on Computer Support for Collaborative Learning*, pages 177–181, 1995.

[IK92] Hiroshi Ishii and Minoru Kobayashi. Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *CHI '92: Human Factors in Computing Systems*, pages 525–532, 1992.

[IM91] H. Ishii and N. Miyake. Toward an open shared workspace: computer and video fusion approach of teamworkstation. *Communications of the ACM*, 34(12):36–50, 1991.

[IUK+94a] Kori Inkpen, Rena Upitis, Maria Klawe, Joan Lawry, Ann Anderson, Mutindi Ndunda, Kamran Sedighian, Steve Leroux, and David Hsu. "We Have Never-Forgetful Flowers in Our Garden:" Girls' Responses to Electronic Games. *Journal of Computers in Math and Science Teaching*, 13(4):383–403, 1994.

[IUK+94b] Kori M. Inkpen, Rena Upitis, Maria Klawe, Joan Lawry, Ann Anderson, Mutindi Ndunda, Kamran Sedighian, Steve Leroux, and David Hsu. "We have never-forgetful flowers in our garden:" Girls' responses to electronic games. *Journal of Computers in Math and Science Teaching*, 13(4):383–403, 1994.

[JJ74] Roger T. Johnson and David W. Johnson. Instructional goal structure: Cooperative, competitive, or individualistic. *Review of Educational Research*, 44(2):213–240, 1974.

[JJ83] Roger T. Johnson and David W. Johnson. Social interdependence and perceived academic and personal support in the classroom. *Journal of Social Psychology*, 120:77–82, 1983.

[JJ86] David W. Johnson and Roger T. Johnson. Computer-assisted cooperative learning. *Educational Technology*, 26:12–18, 1986.

[JJS86] Roger T. Johnson, David W. Johnson, and Mary Beth Stanne. Comparison of computer-assisted cooperative, competitive, and individualistic learning. *American Educational Research Journal*, 23(3):382–392, 1986.

[JJSR85] Roger T. Johnson, David W. Johnson, Linda E. Scott, and Beverly A. Ramolae. Effects of single-sex and mixed-sex cooperative interaction on science achievement and attitudes and cross-handicap and cross-sex relationships. *Journal of Research in Science Teaching*, 22(3):207–220, 1985.

[JMJ+81] David W. Johnson, Geoffrey Maruyana, Roger Johnson, Deborah Nelson, and Linda Skon. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychology Bulletin*, 89(1):47–62, 1981.

[KBW83] James A. Kulik, Robert L. Bangert, and George W. Williams. Effects of computer-based teaching on secondary school students. *Journal of Educational Psychology*, 75(1):19–26, 1983.

[KP95] Maria Klawe and Eileen Phillips. Electronic games engage children as researchers. In *CSCL '95*, Bloomington, IN, October 1995.

[KR93] K. Kelly and H. Rheingold. The dragon ate my homework. *Wired*, July:69–73, 1993.

[Les86] (Main author)Jack W. Lesage. *Journeys in Math 7*. Ginn and Company, Canada, 1986.

[Lev92] James A. Levin. Education on the networks of today and tomorrow: The integration of computer-supported cooperative learning and work. *SIGCUE Outlook*, 21(3):8–9, 1992.

[Mal82] Thomas W. Malone. Heuristics for designing enjoyable user interfaces: lessons from computer games. In *Human Factors in Computer Systems*, pages 63–68, Gaithersburg, MA, 1982.

[Mal86] Mary Male. Cooperative learning for effective mainstreaming. *The Computing Teacher*, 14(1):35–37, 1986.

[McG96] Joanna McGrenere. Design: Educational electronic multi-player games a literature review. Technical Report 96–12, Department of Computer Science, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, 1996.

[MF91] C. Morningstar and R.F. Farmer. The lessons of Lucasfilm's habitat. In M. Benedikt, editor, *Cyberspace: First Steps*, pages 273–302. MIT Press, 1991.

[Mic77] James W. Micheals. Classroom reward structures and academic performance. *Review of Educational Research*, 47(1):87–98, 1977.

[NC93] Bonnie K. Nastasi and Douglas H. Clements. Motivational and social outcomes of cooperative computer education environments. *Journal of Computing in Childhood Education*, 4(1):15–43, 1993.

[OKdV96] R. Oosterholt, M. Kusano, and G. de Vries. Interaction design and human factors support in the development of a personal communicator for children. In *CHI '96: Human Factors in Computing Systems*, pages 450–457, Vancouver, BC, 1996.

[OL93] Sierra On-Line. *The Even More Incredible Machine*. Computer Software, 1993.

[Pap93] Seymour Papert. *The Children's Machine: Rethinking School in the Age of the Computer*. BasicBooks, New York, NY, 1993.

[Per94] J. Pereira. Video games help boys get a head start. *Wall Street Journal*, March 16:B1–B4, 1994.

[Pro92] Eugene F. Provenzo. The video generation. *The American Schoolboard Journal*, March:29–32, 1992.

[Rep93] Judi Repman. Collaborative, computer-based learning: Cognitive and affective outcomes. *Journal of Educational Computing Research*, 9(2):149–163, 1993.

[Res92] Mitchel Resnick. Collaboration in simulated worlds: Learning through and about collaboration. *SIGCUE Outlook*, 21(3):36–38, 1992.

[Rie92] Margaret Riel. Cooperative learning through telecommunications. *SIGCUE Outlook*, 21(3):14–17, 1992.

[RMWW92] J.M. Randel, B.A. Morris, C.D. Wetzel, and B.V. Whitehill. The effectiveness of games for educational purposes: A review of recent research. *Simulation and Gaming*, 23(3):261–276, 1992.

[Rob94] Jenifer Walsh Robertson. Usability and children's software: A user-centered design methodology. *Journal of Computing in Childhood Education*, 5(3/4):257–271, 1994.

[Ros92] Jeremy Roschelle. What should collaborative technology be? A perspective from dewey and situated learning. *SIGCUE Outlook*, 21(3):39–42, 1992.

[SBM+89] M. Scardamalia, C. Bereiter, R. S. McLean, J. Swallow, and E. Woodruff. Computer supported intentional learning environments. *Journal of Educational Computing Research*, 5:51–68, 1989.

[Sed97] Kamran Sedighian. Interface style, flow, and reflective cognition: Issues in designing interactive multimedia mathematics learning environments for children. Doctoral dissertation, Dept of Computer Science, University of British Columbia, 1997.

[Sha80] Shlomo Sharan. Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research*, 50(2):241–271, 1980.

[SK96a] Kamran Sedighian and Maria Klawe. Super tangrams: A child-centered approach to designing a computer supported mathematics learning environment. In *International Conference on Learning Sciences '96*, pages 490–495, Evanston, IL, 1996.

[SK96b] Kamran Sedighian and Maria M. Klawe. An interface strategy for promoting reflective cognition in children. In *CHI '96: ACM Conference on Human Factors in Computing Systems*, pages 177–178, Vancouver, BC, 1996.

[Sla80] Robert E. Slavin. Cooperative learning. *Review of Educational Research*, 50(2):315–342, 1980.

[Sla90] Robert E. Slavin. *Cooperative Learning: Theory, Research, and Practice*. Prentice Hall, Englewood Cliffs, NJ, 1990.

[SM94] Karl E. Steiner and Thomas G. Moher. A comparison of verbal interaction in literal and virtual shared learning environments. In *CHI '94: Human Factors in Computing Systems*, pages 97–98, Boston, MA, 1994. Conference Companion.

[Sol91] Elliot Soloway. How the nintendo generation learns. *Communications of the ACM*, 34(9):23–26, 1991.

[Sol93] Elliot Soloway. Interactive learning environments: Where they've come from & where they're going. Tutorial Notes, Hypertext '93, Seattle, WA, 1993.

[ST79] Robert E. Slavin and Allen M. Tanner. Effects of cooperative reward structures and individual accountability on productivity and learning. *The Journal of Educational Research*, 72(5):294–298, 1979.

[SU97] S. Saxton and R. Upitis. Players, resistors and creators: Adolescents and computer games. Technical report, 1997.

[SW87] S.B. Silvern and P.A. Williamson. The effects of video game play on young children, agression, fantasy and pro-social behaviour. *Journal of Applied Developmental Psychology*, 8:453–462, 1987.

[TM94] P. Thomas and R. Macredie. Games and the design of human-computer interfaces. *ETTI*, 32(2):134–142, 1994.

[Uni97] Simon Fraser University. Virtual-U research project. *http://virtual-u.cs.sfu.ca/vuweb/*, 1997.

[Web82a] N. M. Webb. Peer interaction and learning in cooperative small groups. *Journal of Educational Psychology*, 74(5):642–655, 1982.

[Web82b] Noreen M. Webb. Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology*, 74(4):475–484, 1982.

[Web82c] Noreen M. Webb. Student interaction and learning in small groups. *Review of Educational Research*, 52(3):421–445, 1982.

[Web84] Noreen M. Webb. Microcomputer learning in small groups: Cognitive requirements and group processes. *Journal of Educational Psychology*, 76(6):1076–1088, 1984.

[WM94] J.H. Wiebe and N.J. Martin. The impact of a computer-based adventure game on achievement and attitude in geography. *Journal of Computing in Childhood Education*, 5(1):61–71, 1994.