Evidence from a Series of Experiments on Video-mediated Collaboration: Does Eye Contact Matter?

Richard Joiner  
Department of Psychology, University of Bath, UK.  
r.joiner@bath.ac.uk

Eileen Scanlon  
Institute of Educational Technology, Open University, UK.  
e.scanlon@open.ac.uk

Tim O’Shea  
Birkbeck College, University of London, UK.

Randall B. Smith  
Sun Microsystems Laboratory, California, USA.

Canan Blake  
Institute of Educational Technology, Open University, UK.

ABSTRACT

We report a series of studies on the role of eye contact in video-mediated communication. These are part of an ongoing research program, which is investigating the usefulness of the technological mediated collaborative problem solving for distance learning. Technological mediation consists of access to shared simulations and access to a variety of means of communication. The means of communication we have explored range from audio only contact to video mediated communication with or without eye contact. The motivating question behind this research program is ‘what is different when members of a problem-solving group are physically separated but technologically connected?’ The studies are set in the context of pairs of adults working with shared simulations of either physics or statistical experiments. The first set of studies investigated pairs working on a shared simulation of a physics experiment developed in SharedARK. The study compared remote technological mediated communication with communication that occurred during physical co-presence. There were no differences in performance, but the addition of computer mediated communication did influence the pattern of interaction. These experiments suggested that eye contact influenced problem solving. The second set of studies compared pairs and groups of threes and fours using a simulation of a statistic experiment developed in a system called Kansas. In these studies we compared learners with either video-mediated communication or audio only communication. The addition of the visual communication channel altered the pattern of interaction. The most recent study presents evidence that suggests eye contact facilitated conceptual understanding.

Keywords  
Collaboration, Science, Simulation, and Evaluation

INTRODUCTION

In this paper, we discuss the role of eye contact in technologically mediated collaboration. This is discussed in the context of a ten-year series of experiments designed to explore the usefulness of the technological mediation of collaborative problem solving. The components of the technological mediation are access to shared simulations and access to a variety of means of communication with co-learners. These means of communication range from audio contact only to video conferencing with and without full eye contact. The purpose of this work is the desire to make the future use of such technologies by distance learners as effective as possible. Over the period in which these experiments have been run, there have been developments in the extent to which such technologies are available and affordable for schools and colleges.

This paper therefore reviews this work and as a consequence describes how the use of information communication technology alters the experience of learners in these settings. This programme builds on past work evaluating science based computer assisted learning software with both children in schools and adults in other settings (see e.g. Smith et al., 1991,
Scanlon et al., 1993, Taylor et al., 1993). The focus of the programme is to explore the potential of collaborating technologies for distance learning. The motivating question for the research programme is 'what is different when members of a problem solving group are physically separated then reconnected via this type of computer and communications technology?'

**TASKS MEDIATED BY TECHNOLOGY**

Introducing technology into a setting can have both predictable and unpredictable effects. It is our experience that using technologically mediated collaboration changes the nature of the activity in ways we had not predicted. We discuss these later in the case studies but first we review findings from other empirical studies of technologically mediated problem solving.

Early laboratory studies compared video-mediated communication with audio only communication on cognitive problem solving tasks and found no benefits of video mediated communication in terms of performance efficiency or quality (e.g. Chapanis 1975; Chapanis et al., 1972, Reid, 1977, Short et al, 1976, Williams, 1977). More recent studies have reported similar findings. Fish et al., (1992) compared the use of videophones with audio only telephones and found no difference in usage statistics. Kraut et al., (1996) compared participants working with an expert on a bicycle repair task, with video-mediated communication and with audio only and found no difference in the time taken or quality of the performance. In a follow up study Fussell et al., (2000) compared participants who were located side-by-side or connected via audio-video or audio-only links. They replicated their finding from the previous study and found no difference between audio-only and audio-video conditions in terms of the interaction or the performance of the participants in the audio Anderson et al., (1997) used a map task and found no differences in performance between those used video mediated communication and those using audio only communication.

Only in tasks involving conflict resolution and negotiation is there positive evidence for the benefits of video-mediated communication. Olson et al., (1995, 1997) reported that face to face or video-mediated communication produced higher quality designs than groups who used only audio only communication. Veinott et al, (1999) compared the performance of people explaining a map route to each other. Half were native speakers of English and half were non-native speakers of English. They found that non-native speakers benefited from video-mediated communication whereas there was no difference in performance of the native speakers who used video-mediated communication compared to those who used audio only communication. Thus research in this area seem to suggest that the benefits of video mediated communication are be dependent on the nature of the task.

Although commercial video conferencing systems are becoming more prevalent, there are a variety of parameters which require further examination to establish fitness for purpose. O’Malley et al., (2001) report on work in the Eye-2-Eye project which aims to provide comparative data in order to understand the different uses and cost benefits of different technologies. One of their interesting findings from this project is the decreased likelihood of lying when pairs were able to see each other using a videoconference compared to when they used an audio-only conference. Fulwood (2001) has investigated the problems in communication associated with the positioning of the camera in relation to the monitor. In this set-up direct eye contact is impossible and the lack of it was reported a problem by the users. His solution was to train them to gaze into the camera rather than the monitor.

**CASE STUDIES OF SYNCHRONOUS COLLABORATION**

We have been studying synchronous collaboration in the context of adults working on shared simulations. The main technological base has been provided by a number of systems designed by Randall Smith. The Alternate Reality Kit, ARK (Smith, 1992) was a system for creating interactive animated simulations implemented in the Smalltalk-80 programming environment. The Shared Alternate Reality Kit, SharedARK, (Smith et al., 1991) and KANSAS, a network shared application space (Smith, in press) are prototype technologies for allowing students to work together at a distance from each other on a shared simulation while maintaining voice and eye contact. The following are two simulations used to explore the capabilities of the shared space for problem solving, the Running in the rain experiment and the Gameshow experiment.

**The ‘Running in the Rain’ experiment**

Our primary purpose in this experiment was to assess the usability of the SharedARK technology and to identify factors which were important in facilitating collaborative problem solving with this technology. We did this by comparing remote electronically mediated communication during collaborative problem solving, with that occurring during physical co-presence. These experiments involved eleven pairs of adults working at a simulation of an under-specified problem. Four of the pairs of participants used the video tunnel while working together on the simulation, three worked on the simulation as co-present participant pairs and two pairs worked remotely with only audio contact. We used the problem of whether to run or walk in the rain without an umbrella (De Angelis, 1987). Running means spending less time in the rain, but, on the other hand, since you are running into some rain, you might end up wetter than if you had walked.
In these experiments two users are in separate rooms with a workstation each, and communicate through a high fidelity, hands free audio link and with a camera/monitor device called a video-tunnel which enables both voice and eye contact through the use of a beam splitter and a mirror (see figure 1).

![Video Tunnel Diagram]

Figure 1: Video Tunnel

After an eight to ten minute introduction to the interface and the task, the participants were given a simulation containing a rain cloud, a rain runner and a device to control simulation parameters such as the speed and the direction of the rain and the wetness of the runner. During the introduction they are told that the object of the activity is for them to jointly agree when it is worth running in the rain. They are shown how to make the runner wider and narrower, how to make the runner move and how to switch the rain on and off. Participants were invited to use the simulation to test their ideas and after ninety minutes they were asked to report on their findings.

A data capture suite was used to capture video records of the interaction, which were then displayed in a four-way matrix. For pairs this meant capturing each user’s screen and their video conferencing record. (For the larger groups a single shared screen was recorded with each individual user’s video.) Video cameras were used to record task performance. One camera was used to record task performance. One camera captures the information displayed on screen. This data collection set-up was inspired by the development of the media space at Xerox PARC in the mid-eighties (see Bly et al., 1993), and then developed further in Rank Xerox research centre Cambridge where the running in the rain experiments were recorded. This facility was then developed in the Computers and Learning Research Group at the Open University where later experiments took place.

Four synchronous video signals made up of what was displayed on the two video-tunnel screens and the two simulation screens worked on by the participants. The video protocol was analysed by relating utterances made by participants to both events in working with the simulation and eye contact. Participants' activities and utterances from their verbal protocols were categorised according the type of activity in which participants were engaged, according to whether it involved the interface, (e.g. figuring out how to alter the rain runner's speed, or the rain runner thicker), the task (discussion about running in the rain or social interaction (e.g. laughing at jokes). These utterances were then further assigned to subcategories called meta-level activity, specific activity and recovery, which further describe the nature of the activity. For example, meta-level activity might be generating hypotheses, or discussing problem-solving strategies, specific activity might be talk generated in doing the task, and recovery from breakdown might be recovery from interface errors or misunderstandings during a conversation. Features of the dialogue are described in Taylor et al. (1993), and the type of augmented problem solving which was facilitated described in Scanlon et al. (1993), and the nature of the shared space and its role in establishing successful collaboration described in Smith et al. (1991).

This analysis revealed some interesting features. Use of the video channel was correlated with activity when participants were not directly manipulating the interface. We found several examples of jokes being cracked across the video channel. There are long periods where the video tunnel was not used at all, for example when objects are being manipulated or data points collected. However when the participants were talking about what they observe or suggesting hypotheses or planning experiments they look towards their partner through the tunnel. Another difference was that participants who could see each other were terser and less explicit. We hypothesised that a video channel might encourage non-interface specific activity and indeed meta-level discourse about the task was accompanied by much higher levels of eye contact than was specific talk about the interface. There was also considerable differences in the way that different pairs negotiated their problem solving.
The unique shared space created by the participants using the technology is interesting. The use of overlapping (shared) and non-overlapping work areas seemed to have encouraged task division but also role division was fluid. Participants can both be face to face through the video tunnel and side by side looking at the SharedARK interface. This would suggest that pairs using the video tunnel might do better in solving the problem since being face to face facilitates patterns of mutual gaze. We thought that more productive problem solving might ensue. In considering the use that participants made of the video tunnel, which coincided with joint assent to particular decisions particularly at the planning phase. This side to side and face to face combination was not available to co-present pairs, and seemed to influence how role and task division takes place. In the audio-only condition there was more explicit negotiation of task division than in the video condition.

In substance it appeared from this exploratory experiment that the addition of a video channel influenced the users activity by encouraging interaction about the problem. Comparing the problem solutions produced by pairs on this task with written problem solutions (see Scanlon et al., 1993) the study demonstrated that the use of audio links and video tunnels does not attenuate the problem solving behaviour of pairs working on the task but in fact enhances it. In particular this experiment highlighted the importance of eye contact in joint problem solving.

The GAMESHOW experiment
This work informed the design of a second series of studies exploring a statistics-based simulation implemented in a distributed classroom environment designed by Randall Smith at Sun Microsystems called KANSAS. This allows several physically distributed users to work together in a shared virtual environment. Moving together and apart in Kansas will make and break audio connections. Users each have a window in which they can see their local portion of Kansas. Each user sees a small local rectangular portion of Kansas and can scroll their viewpoints across the vast surface, causing their rectangles to overlap in order to collaborate, or can move away from others to work alone. Our experiment here was conducted on a version of Kansas which supports up to 5 simultaneously active users (or, of course, up to 5 groups of users) and users can be given access to audio or both video and audio links between each of the 5 locations. Randall Smith has experimented with larger numbers of simultaneous users (see Sipusic et al., 1999 for another educational use of this technology).

We used a game show simulation of a well-known statistics problem- the Monty Hall dilemma to explore a number of related themes.
• The effect of working with a simulation on concept development in statistics.
• The influence of the bandwidth of the communication channels on the collaboration.
• The scalability of the collaborative experience.
• The usability of the interface design.

We have collected data on 6 pairs of participants, two groups of three and a group of four using the shared simulation augmented by audio communication. We also have collected data on five pairs who used a video tunnel to provide a video channel of communication, and a further four pairs working with altered video tunnels where there is no eye contact. In this paper, we are comparing data from pairs working with the video tunnels and pairs working with the altered video tunnels.

We are using a simulation of a well known statistics problem- the Monty Hall dilemma originating in an TV gameshow called ‘Let’s make a deal’ where the gameshow host encouraged contestants to make a choice between three items, and then change their choice. This problem is non trivial and caused extensive correspondence between statisticians when discussed in a newspaper (see Hoffman 1999). The groups of participants were asked to explore the problem with the aid of a shared simulated game show setting, a shared note-taking tool and a remote human host. They communicate over an audio or video link. The game show host displays the consequences of their choices. We have recorded videos of adults at working together on the statistics simulation in different physical locations and observed their problem solving behaviours and the impact of the experience on their understanding of concepts in probability.

The participants were interviewed and given an individual pre-test questionnaire. The participants were told.
‘You are a game show contestant. You have won through to the final round and your final challenge is to choose one of three doors. Behind one but only one of the doors is a Mercedes. You announce your selection but before you open the door the game show host ‘helpfully’ opens one of the doors which was not the one you have chosen. It doesn't have a car behind it. What should you do, stick to your original choice or change’

In the pre-test, they were asked individually to make a prediction and to give a reason for that prediction. Then they were introduced to their partners and given a shared simulation to conduct experiments. The time taken on the task ranged from 30 minutes to 90 minutes. After the simulated experiment, students still in their groups were asked to make a joint statement of their solution and to comment on whether it had changed from their individual statements. Then they completed an individual post test questionnaire to establish what their own opinion was. They were asked to state what they
thought the best strategy for the game contestant to pursue. They also were asked to make a prediction about what the best strategy to pursue if the problem had four doors and the game show host opened two of them.

When groups devised simple highly focused ways of summarising the data that could be viewed in a single window they were much more likely to identify and resolve their mathematical misconceptions. The most elegant summary was a two by two matrix of success/fail versus stick/change where the appropriate cell was updated by one of the participants after each go at the game.

In this experimental situation we describe above, the learning experience was managed in part by the game show host and this raises another issue related to scalability- the issue of learner support. In this case the learning support demands appear to be much more acute than those found in the conventional classroom (see O'Shea, Smith, & Scanlon, 1997).

This system was designed to allow participants working collaboratively to solve particular problem. The particular focus here was on trying to understand how students could use a system which allowed them to conduct variable based practical experiments to help them develop their knowledge and understanding of a statistics topic. Simulations on computers can allow many experiments to be conducted quickly to develop an understanding of statistical topics. We found some success with this method (the majority of students made progress in their understanding of the problem in addition to the detection of many of the misconceptions about randomness cited above. Students however held widely differing views about how many trials of different strategies were necessary to build up a sensible picture of the outcomes of different strategies. A full account of their attempts is given in Scanlon et al. (1997).

As to the influence of the means of communication on collaborative problem solving, we were able to compare the behaviour of pairs communicating with an audio channel only or a video channel. The influence of the video channel was less marked than in the previous Shared Ark experiment in terms of pattern of discourse. Some pairs using audio were fairly terse, but some were quite discursive. The use of audio only however did require more interchanges clarifying task division. For example one pair communicating over audio only about their use of the shared note taking tool comment:

X Sorry, D. am I blitzing you?
Y Yeah I didn’t realise that, so who is going to… what shall we do about... shall we have a procedure for making the notes so we both don’t try and type at the same time?
X OK
Y Shall I type in?
X You go ahead and type in yes

However, in some pairs, the access to and use of the video channel did lead to a better co-ordination of views between the two participants, what they thought the position was what they thought a good experiment to do was. One pair in particular used the video channel to explain their current view of the problem was even drawing diagrams to explain what they thought the explanation for the successful strategy was. In addition this pair’s conversations over the video channel had a particular courteous quality

X Do you want to do any more, or do you think we’ve..?
Y I’d quite like to try ten more with not changing
X OK You’re best at writing, do you think, it doesn’t depend on who chooses does it?
Y I don’t think so
X It’s the policies what count, so do you want to choose and I’ll count this time?

We have also begun to analyse the effects of technological mediation on conceptual change. We compared the pairs using the video tunnel with the pairs using video-mediated communication without eye contact: in terms of the number of participants who changed from saying stick was the best strategy to saying that swap was the best strategy. Table 1 shows the results from this comparison.
Interestingly the pairs using the video tunnel were significantly more likely to change their opinion than participants using video mediated communication with no eye contact. The audio only and the VMC without conditions were collapsed together because in the table above more than 4 cells had an expected frequency less than 5 (Chi-squared = 5.5, df = 1, p < 0.05). This tendency to change made them more successful in their problem solving.

### Table 1: Number of participants who change their opinion.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Change</th>
<th>No Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio Only</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>VMC with Eye Contact</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>VMC without Eye Contact</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

COMPARISON OF THE CASE STUDIES
The two systems described here differ in the extent to which the computer-based component and the audio-video component are distinct. In Shared Ark they are disjoint, as the audio/video is output to a device completely separated from the computer screen, while in KANSAS the audio/video is integrated with the interface. However more fundamentally for both systems the video and computational components are separate. For the use of the systems that we have explored this is not an issue, as we have used audio/video only for communication between learners but the possibilities of integration between video and simulation lead to possibilities for more extended experimentation.

Videoconferencing is an appealing possibility for distance education. Multiple video images can be displayed on a single monitor (for effects see Smith, in press). For pairs of students the augmented video contact of the video tunnel provided a sense of eye to eye contact. Multiple video sources can be displayed together but there is no way yet to simulate this improved eye contact for larger groups of users.

Comparing the two simulations they had some similar features but also some differences. For example, the Running in the rain simulation required more attention from participants, while the Game show simulation was more dependent on the successful use of the shared note-taking tool. In addition, in both experiments, as well as acting as an experimental resource, the shared notebook implemented on the screen became a shared focus for discussion (see e.g. Enyedy et al., 1997).

CONCLUSIONS AND FURTHER WORK
The motivating question for the research described here was to explore situations when members of a problem solving group are physically separated then reconnected via combinations of computer and communications technology. The provision of a simulation which members of the group could explore is a key feature of the systems we describe here. Our finding in general is that distributed problem solving can be supported by appropriate technologies without attenuation, but we have also noticed ways in which the interaction is altered by the technology. If we consider a pair of individuals working together in the same room, the way in which they interact will vary over time, their proximity will vary. When a technology is the medium of communication between two physically separated individuals the proximity relations will be altered. They will not be able to touch but equally they cannot remove themselves from the interaction without breaking it off altogether. They can work side by side while in the video condition be face to face with their partner. So although there may be a loss of quality in the communication via audio or video channels the proximity relations may be enhanced with an effect on talk about the problem. (See Smith et al., 1991 for further discussion of this.)

The key dimensions of variation which have been explored in these experiments are the number of learners working together, whether or not they are physically co-located, and the bandwidth of the communication channels available to them. The bulk of experience reported here is with pairs of participants communicating remotely over audio or video with physically co-located pairs used to draw comparisons. Indeed the main result of this work is to develop a picture of the shared space created by shared simulations and video communication tools. In both cases, the participants quickly learn to use the simulations and come to terms with the shared audio-video-computer space in which they find themselves. The shared space created by this technology places participants into a kind of enhanced proximity in which it is possible to be simultaneously side-by-side and face-to-face. There is much still to be explored about how such workspaces can be designed to maximise the beneficial effects of collaborative problem solving. We can in these case studies demonstrate particular ways in which participants have used this rich shared resource to augment and facilitate their joint problem solving, which gives us considerable hope that such systems will be developed for distance learners. We find the role of eye contact in video-mediated communication is important in both our first and second set of studies.
Sometimes it has been claimed that the nature of the activity change caused by the use of technologically mediated collaboration alters the authenticity or reality of the learner’s experience, either positively or negatively. However the change in the working practices of scientists over the last few years while we have been engaged in this research means that communication over networks as part of collaborative working is a now common part of the working practices of modern scientists. Therefore the context of this work which began as a laboratory investigation of an unfamiliar and futuristic setting for group working is one which today’s students now accept as a reflection of the type of settings they may encounter in their future working lives. Students and teachers can confidently predict the continuing integration of information and communication technologies with other tools and educational researchers can expect to build and test virtual learning environments, focussing on identifying the different special properties of such new shared spaces.

ACKNOWLEDGEMENTS
The authors gratefully acknowledge the members of the Running in the Rain project team (Josie Taylor and Claire O’Malley) and Game show project team (Yibing Li, David Perry, Mark Treglown, Canan Tosunoglu).

REFERENCES


