

The Virtual Tea Room – Experiences with a New Type of Social Space.

Renee Gedge

School of Computer Science & Software Engineering,
Monash University, Melbourne, Australia
+61 3 99032572

renee.gedge@infotech.monash.edu.au

David Abramson

School of Computer Science & Software Engineering,
Monash University, Melbourne, Australia
+61 3 99032572

david.abramson@infotech.monash.edu.au

Abstract

This paper focuses on technical, social and psychological aspects of the use of a video wall for communications between two geographically separated halves of a university department. The overall usability of the system is related to technological limitations such as video and audio quality, and aspects connected with organizational and interpersonal psychology. A video wall creates a new type of “social space”, raising new social psychology issues, and requiring new types of interpersonal behaviour. Whilst it has some close parallels with standard desk top video conferencing, it also has some unique and important differences. This paper briefly summarizes those problems and discusses the issues that are shared with other electronic collaboration tools such as desktop video conferencing, followed by a discussion of those issues unique to the video wall. Ongoing trials are being conducted with alterations to audio and video quality, and an analysis is being performed of usage of and responses to the video wall.

1. Introduction

In late 1999, The Monash University School of Computer Science and Software Engineering set up a *Virtual Tea Room*, an environment which is physically two large rooms on two campuses 8 kilometers apart, with a video wall in each room. Photographs and brief history can be viewed at the Virtual Tea Room website [1]. The system uses MPEG-2 CODEC-based technology from FVC.com, allowing the transmission of life-sized high quality video and audio data over the University's ATM network. The initial purpose of the video wall was twofold – to allow face to face meetings between members of the school as a group, and to provide staff with opportunities for casual social interactions across campuses. The Virtual Tea Room/Video Wall has now been in place for a full year, and a number of observations can be made about the experience.

2. Electronic communication and human factors

Considerable research has been conducted into negative social effects of electronic technologies such as email, and tele/video conferencing. Human interactions are both complex and subtle, and may be adversely affected when this richness is reduced. Visual cues such as body language and eye contact, the need for conversational fluidity and topic flexibility, and respect for existing conversational conventions are all important aspects of human communication. In general, electronic communication systems are deficient in supporting one or more of these characteristics.

It has been established [2,3] that there are significant differences between video conference systems communication and face to face communication. One of the most basic is that there is less of it – in particular, informal communication is much less likely to occur than in a face to face situation. [2]. Users report feeling constrained by several aspects of the system. One of the most commonly reported sources of unease is the lack of reciprocity and privacy inherent in the system. Reciprocity is the sense that if you can see and hear someone, then they can see and hear you. There are exceptions to reciprocity in the physical world - such as someone who is invisible but watching you, a telephone connection that is only audible one way, or even talking to a sight or hearing impaired person. These situations similarly cause us unease or difficulty. Some videoconferencing systems have tried to address this issue by ensuring reciprocity to the extent that either both ends are working or neither is. However, this does not reassure all users. It is often not even apparent if this “rule” is being applied, and it does not solve the problem of participants who may be listening and watching, but are out of camera range. Kraut [4] reported that new users were commonly apprehensive about using the videoconference system as they were concerned they may have been interrupting someone else – ie someone they could not see or hear to be speaking then. It is a basic conversational convention not to interrupt someone else, but support for this convention within videoconferencing systems is more limited than face to face interaction.

Lack of reciprocity raises the issue of privacy with most electronic communication as it is impossible to tell who else may be watching and listening from off camera. In addition, it is impossible to use traditional means of ensuring privacy with another person such as moving closer and speaking quietly.

Users have also reported that limited camera fields of view, unnatural camera angles, and the need to take account of microphone positions all tended to inhibit the natural flow of communication [2]. It is certainly true that most current systems impair conversational fluidity.

Fish et al [2,3] also reported that the absence of some important mediators of communication, such as eye contact, increased people's sense of unease with video conferencing systems. Even with high quality links, it is easier to avoid eye contact and at worst – depending on image quality – impossible to maintain it even with the best intentions. On the other hand, people may welcome the opportunity to communicate with reduced social interaction where this means the avoidance of an undesirably open or intense interaction. In a case study of email use, Markus [5] found that some users deliberately chose to use email for communication because it allowed them to avoid unwanted social interactions. Marcus found that the social and interpersonal impact of email technology derived from both intentional and entirely unintentional outcomes of the use of the technology. Using email to limit interaction was a deliberate outcome. However, at other times, email communications did unintentionally have a negative impact on relationships between users owing to limitations in expressiveness of the medium. This aspect has been used to explain the prevalence of so-called "flame wars" in which exchanges of email escalate to progressively more heated and abusive levels, the content being interpreted without the benefit of visual or auditory cues, interactivity or other modifying factors.

A number of studies comparing face to face versus video conference supported group collaboration have found that even when users are quite satisfied with quality of solutions and learning, users still expressed dissatisfaction with group interaction processes and quality of discussions [7, 9]

It was in attempting to determine and describe exactly what is missing from what appear to be sophisticated electronic communication systems that Erikson & Kellogg [6] described the concept of "social translucence". They argue for the design of electronic communication systems that exhibit three primary characteristics - those of visibility, awareness and accountability. These characteristics allow the transmission of socially important information and support the observation of social norms and rules which people normally use to govern interactions. Whilst these characteristics usually occur together in the physical world, they do not necessarily occur together in the digital

world. A simple example of *visibility* in communication is the visual transmission of all relevant human communication data – this includes detailed facial expressions, body language, tone of voice etc. Many of this data is entirely missing or greatly reduced in videoconferencing communications.

Erikson & Kellogg also point to the importance of a "shared awareness of constraints" amongst participants in a communication. For example, if everyone knows there is a lot of background noise in an environment, then they will make adjustments to compensate, such as moving closer or raising their voice. Not only will they do this, but they will know they share the awareness of this need with everyone else present, and will expect similar behaviour from others present. It is an integral part of the communication set up – an assumption of commonality. This also breaks down in many digital communications, of which the video conferencing situation is a good illustration. At one end, sound and picture may be excellent, whilst at the other one or both may be poor or absent without the other participants being aware of this. Therefore, participants may not only lack important social information, but they also cannot rely on shared awareness of any existing deficits affecting the communication.

With regard to the notion of accountability in communications, Erikson & Kellogg contend that accountability for our actions is a very important component in human communications, with significant negative consequences if it is missing. If people in a videoconference ignore asides, signs of distress or disagreement from "the other side" they may not be held accountable for what would otherwise be seen as rude or insensitive behaviour, as no one can be sure they are even aware of them.

These aspects of human communication, visibility, awareness and accountability, are of varying importance depending on the communication situation. The following section uses these concepts to examine experiences with the video wall in both formal and informal communications.

3. Experiences with the video wall.

The video wall is used for a number of quite different functions – including formal monthly school meetings, seminars, and ad hoc smaller meetings between staff members. In addition to these are spontaneous social interactions between anyone in the virtual tea rooms, which was always a primary purpose of the system. The rooms actually only support one of the basic communication options well, specifically one to many (1:M). In its current form it cannot support one to one (1:1), except when there happens to be only one person at either end, as the speakers' voice will be heard by anyone

else in the room at the other end. Many to many communication (M:M) only works well in small groups, and does not scale well to larger groups. This is because the system currently relies on a speaker's proximity to a microphone, with this audio being broadcast at the other end. In general, to be effective, many to many requires broadcasting input from each microphone to different zones within the room. Whilst these technical problems could be addressed by stereo or multi-channel transmission, the current configuration using single channel sound limits the nature of communications possible. In spite of this, we have witnessed small groups undertaking multiple independent conversations concurrently through the wall. In these cases, it seems that the addition of a video of someone taking is sufficient for a person to disambiguate the multiple conversations. If the video was not present, then this would be very difficult.

4. Video wall use for formal meetings.

School meetings generally involve some 15-20 staff at each campus and are quite formal in nature, with a chairperson and minute taker. School staff generally agree that attendance at these meetings has been facilitated by the video wall. For example, prior to the installation of the wall, monthly staff meetings alternated between campuses. Staff from the "non-home" campus were invited to attend, but rarely did. Thus, the full staff meetings defaulted to campus based meetings. After installation of the video wall, we found that roughly equal numbers attended each monthly meeting. The technology allows staff to attend meetings without the disruption and wasted time associated with travelling the 8 kilometres between campuses. In spite of this, there are mixed feelings about its impact on the quality of interactions at these meetings. Some staff felt that the communication was compromised by the quality of video and audio. Some specific and common complaints were that the video quality did not allow for the interpretation of facial expressions and body language. In particular, only reduced eye contact is possible with people at the other end, even when two participants are both sitting quite close to the camera. However, eye contact would also be somewhat reduced within such a large group in a face to face meeting. In spite of these problems, other staff feel the functionality offered by the system is so *enabling*, that they forgive the inaccuracy of the medium.

We did not experience problems with latency. Typically the latency between the ends of the wall is less than the .4 seconds which has been previously identified as the critical level [7].

More importantly, audio quality was sometimes inadequate for ready comprehension. This has previously been shown to be critical for video conferencing success ([8,9]. A satisfactory level of communication quality can

be maintained with relatively poor video quality as long as audio is good [8]. Audio problems usually occurred when the person speaking was at some distance from one of the two microphones. Dealing with it requires an interruption from the other end to indicate a problem with audio, and then the microphone or the speaker must be moved, thereby significantly disrupting the flow of discussion.

It has been observed in numerous meetings that people will interrupt to request improved audio a number of times early in the meeting. However, as meetings progressed, staff were observed to reduce these interruptions, apparently accepting they would be unable to hear the discussion. There was sometimes an exchange of frustrated or resigned looks among participants at the receiving end. It seemed there was a reluctance to keep interrupting, a form of embarrassment, rather than a loss of interest in the discussion. This problem often resulted in a prolonged discussion almost entirely confined to one side of the video wall, with only a limited amount of its content audible to the other end, but with the "discussing end" apparently unaware of the problem.

A related aspect of this was the sense amongst active participants in discussions of missing feedback from the other end. The slight nods, eye contact, or murmured assents that often constitute feedback while we are speaking in this context are missing as they are too subtle for the system to transmit adequately. The speaker(s) may be left unsure whether they have been heard, and if so what the response actually is. One author's experience was that it was somewhat unnerving to speak into a kind of void. Clearly the provision of good audio is so important that it requires further investigation. We originally thought that two microphones would be sufficient, but in groups of 15 or more we would require at least 4 to cover the area well. Using multiple microphones introduces an additional complication because they introduce more than one audio path, and this interferes with the echo cancellation hardware. This problem might be best addressed by using multiple audio channels, and is currently under investigation.

Another difficulty was the possible presence of "invisible" participants at both ends - ie people present at the meeting or just in the room who were outside camera range. This occurred either because people actively sought a position where they were outside camera range or they were completely unaware that this was the case. When questioned, about their inclination to take a seat "off camera", most people found it difficult to give a simple explanation. This is in itself significant - that people are dealing with an unfamiliar set of circumstances and they do not have many prior or parallel situations to draw on. There was certainly a sense of "being on television" for many people, even though we strenuously avoided the use of television monitors or even a local feedback monitor. In general, people were largely unable

to expand on their discomfort or how it was different to being present and visible in a normal meeting.

On reflection, it seems possible that there is a sense of a lack of knowledge about how they are appearing and sounding at the other end. An interesting illustration of the "awareness" deficit in the current system was given at its formal launch. A few participants were observed to sit down immediately in front of the camera, and stare or smile into it for prolonged periods of time. From the other end this was perceived as a person sitting and looking/smiling at nobody or nothing in particular for a prolonged period of time - a bit like a child grinning into a camera. The participants seemed to have a sense of being "on camera" and therefore just to be there and grin was appropriate. We have observed that for some individuals this response decreases after a period of familiarization. It also decreases when participants are actively involved in a conversation, and their attention is drawn away from the technology and towards their interaction.

At a "normal" meeting, people make choices about where they sit - for example some people will choose to sit a long way from the chairperson, as close proximity would mean they would frequently share the general visual attention given to the chairperson. The video wall situation reduces the sense of control over how much or little others are seeing of you - unless you deliberately sit out of camera range. Not sitting near the chairperson is no guarantee of reduced exposure. In fact in our meetings the chairperson frequently sat at the top of the U shaped seating in order to appear at the center of the meeting. This however made him the furthest from the camera. It also meant that those furthest from him at the same end would actually be closest to the camera and therefore the most salient (closest and most detailed) people on the screen for those at the other end. This is a reversal of the normal situation and may cause significant psychological discomfort to those who sought to "keep a low profile". A solution to this is for the chairperson to sit in the middle of the "virtual meeting" i.e. at one end of the U-shape, closest to the video wall and the camera.

Considerable efforts were made during the year to reduce the amount of room space outside camera range, in order to reduce staff anxiety about unseen observers. However, it proved difficult to eliminate all such space, and reduction of it involved a tradeoff between wider angle view and reduced image quality. Even had all such space been eliminated the most important aspect may still be lacking - ie some people may still experience anxiety that there may be parts of the room they cannot see. A complete view of the entire room is of course difficult to achieve with a single camera and projection system. An early alteration to the physical setup was to locate the camera in a recessed area behind the wall. This has a three fold effect. First, it minimizes non-visible space. Second, it reduces the participants sense of being "on

camera". Finally, it improves the eye contact because the camera is centred in the image.

5. Video wall as a virtual tea room.

A unique aspect of the video wall has been its continuous presence in a space which has the primary function of promoting relaxation and casual conversations between staff members. Unlike a desk top video conferencing system, or a purpose built video conferencing room, these rooms at both campuses were primarily tea rooms. Staff entering the tea rooms would see whoever was present in the tea room on the other campus.

A most interesting aspect of the "virtual tea room" function of the video wall was the strong negative feelings which rapidly developed towards it amongst a small group of staff. These staff members expressed their dislike of the system's presence and asked for controls to turn it off when they wished to. The primary reason given for these negative feelings was a sense in this group that they may be watched whilst unaware of it. In fact this was a real possibility technically, even if an unlikely one socially. The video wall was on even if a person didn't trigger the motion detector. This would allow someone to stand out of view but watch a person at the other end.

When we have a conversation with someone in a room, we can usually see all the boundaries of that room and perceive whether it is unoccupied or not. However, there are also rooms that may have partitioned areas such as a kitchenette, or simply partial partitions in office areas. People do adapt quite readily to these type of environments, such as avoiding making private phone calls which may be overheard. An important difference however is that if you do get up and walk around it is possible to ascertain exactly who is there and where they are in relation to you. This then allows you to judge whether you can conduct a private conversation. Without additional cameras or other sensor technology, this is not possible in the current configuration of the tea rooms.

Some staff seemed to just dislike the sense of being "on camera" and went to considerable lengths to discover which parts of the room were off camera and they would deliberately sit there for their tea or lunch breaks. Others actually placed objects over the camera and disconnected the microphones to avoid the system. Eventually, some staff requested a system whereby people were given access to switch the system off when they were in the tea room at either end. It was felt by the system organisers that this was defeating the purpose of the virtual tea room, namely a space that was genuinely shared by all members of the school. However, in response to this request a few changes were proposed. First, an individual can veto the system by pressing an off-switch. This turns the system off for a fixed period, after which it reverts to automatic mode. Second, the system is only enabled if motion is

detected at **both** ends. If someone is only present at one end, then it is not possible for them to be viewed from the other end, even if there is someone there who has not triggered the motion detector.

Negative feelings towards the virtual tea room were complicated by the fact that the two campuses were previously two separate departments, and even earlier were actually two separate institutions with a quite different focus and method of operation. Whilst many of the earlier differences and even antagonisms have faded, there is still a strong sense amongst some staff of belonging to one campus and one group on that campus. Thus, some staff are not particularly interested in communicating with "colleagues" on the other campus, and this probably increases their dislike of the system.

Despite these difficulties, the virtual tea room has been used quite extensively by many staff members for impromptu social and work related conversations. We have held welcome and farewell parties, met visitors when the individuals could not be on the same campus and held social events that would otherwise not have been possible. The most significant difficulty in this use is the lack of privacy, which can be both perceived and actual. It appears that for many staff it fits into a continuum of communication facilities - somewhere between a telephone call and a face to face meeting. However, this is not a linear progression from least to most complete communication, as telephone calls whilst lacking body language and eye contact cues, can offer a sense of actual and perceived privacy that is impossible with the video wall.

6. Future possible enhancements

An interesting remark made by a clinical psychologist on entering the virtual tea room for the first time was "Has it come to this?". Upon discussion and reflection, the psychologist clarified the thoughts and feelings which led to this remark. These were based around a sense that this type of communication was a poor second best to face-to-face communication, removing some very valuable aspects of human interaction whilst preserving an outward appearance of being very close to the real thing. Perhaps part of the unease generated by the system is due to people recognising that there are differences and deficits with this form of communication, but being unable to readily pinpoint what they were. It may well be that as people become accustomed to the technology, their sense of unease will be replaced by a knowledge and acceptance of its benefits and deficits. We have certainly witnessed this to some extent in our work. However, significant improvements can still be made which address some of the concerns that have been expressed.

The privacy issue is the major social problem when using a video wall for informal communications. This has been at least partially addressed at Monash by making the

system switch off unless it is triggered at both ends. Invisible space also needs to be minimized. The concept of zones that would allow multiple users to conduct private one-to-one conversations needs to be further explored from a technical viewpoint.

Audio quality is the bottom line as far as useability goes for formal meetings. This can only be addressed by more/ better microphones, multiple audio channels and better echo cancellation technology. It is possible some social techniques can also improve communication. For example, beginning a conversation with a simple "Can you hear me okay?" can prevent a lot of frustration and interruptions. High quality visual data should be maximized as it allows the transmission of socially important information such as eye contact and body language. Where limited bandwidth means this quality is compromised, perhaps new social protocols are necessary to facilitate communication. If subtle clues such as body language, eye contact, sighs or murmurs of discontent may go unheeded, meetings may need a more formal mechanism for eliciting disagreement or discontent. The chair of the meetings has found that it is more difficult to chair a meeting using the video wall in order to avoid small splinter groups breaking away from the main meeting. When this occurs in a "normal" meeting it is annoying, but when it happens in a virtual meeting it can be disastrous. It may also be necessary to place quite strict time constraints on meetings to alleviate the fatigue and the consequent drop in quality of concentration and participation which results when audio and video quality are inconsistent or poor.

The Monash SCSSE is currently undertaking extensive additions and improvements to the video wall system which will result in three separate systems, one dedicated to formal meetings, one for seminars and the other for casual, spontaneous interactions. This will allow each environment to be modified to optimize its usability. Ongoing study will be conducted of the human interactions and reactions to these three systems.

7. References

- [1] <http://www.csse.monash.edu.au/~davida/virtualtearoom/>
- [2] Fish, R.S., Kraut, R.E., & Chalfonte, B.L. The VideoWindow system in informal communication." Proceedings of the conference on Computer-supported cooperative work, October, 1990, Los Angeles, CA USA, pp 1-11.
- [3] Fish, R.S., Kraut, R.E., & Root, R.W. Video as a technology for informal communication" Communications

of the ACM, vol 36, Issue 1, 1993, pp 48-61.

[4] Kraut, R.E., Rice, R.E., Cool, C. & Fish, R.S., Life and Death of new technology: Task Utility and Social influences on the use of a communication medium", Proceedings of the conference on Computer-supported cooperative work, October, 1994, Chapel Hill, NC, USA pp 13-21.

[5] Markus, M.L. "Finding a happy medium: explaining the negative effects of electronic communication on social life at work". ACM Transactions on Information Systems. Volume 12, Issue 2 (1994) pp 119-149

[6] Erikson, T. & Kellogg, W. "Social Translucence: An approach to designing systems that support social processes". ACM transactions on Computer-Human Interaction, Vol 7, No 1, March, 2000. pp 59-83.

[7] Ackerman, M.S., Hindus, D. Mainwaring, S.D, Starr, B. "Hanging on the wire: A field study of an audio-only Media space." ACM transactions on Computer-Human Interaction, Vol 4, No 1, March, 1997. pp 39-66.

[8] Ocker, R.J. & Yaverbaum, G.J (1999). "Asynchronous computer mediated communication versus face-to-face collaboration : results on student learning, quality and satisfaction." Group Decision and negotiation. Vol 8(5), Sep 1999, pp 427-440.

[9] Habash, T.F (1999) The impact of audio- or video-conferencing and group decision tools on group perception and satisfaction in distributed meetings. Psychologist-Manager Journal. Vol 3(2), 1999, pp 211-230.

[10] Warkentin, M., Sayeed, I., & Hightower, R. Virtual teams versus face-to-face teams. An exploratory study of a web-based conference system. in Kendall, E. (Ed) et al (1999) Emerging Information technologies: Improving decisions, cooperation, and infrastructure. pp 241-262 CA, USA: Sage Publications, In

