Videoteleconferencing: Why is it disadvantageous for Group Collaboration?

Phillip Jeffrey
German National Research Center for Information Technology
53754 Sankt Augustin, Germany
E-mail: Phillip.Jeffrey@gmd.de

Teleconferencing is the use of a telecommunication system to electronically connect physically separated people [3]. In collaborative working environments, videoteleconferencing is perceived to be an alternative when face-to-face meetings are not possible. The assumption is that the audio and visual capability can replicate the face-to-face environment of ‘being there’ [3, 13].

Nonverbal behaviour such as gaze and gestures and verbal behaviour such as backchannelling are used as feedback in face-to-face communication to naturally regulate the flow of conversation by providing cues for turn-taking and floor control [1, 8]. According to Sellen [13], one’s head and shoulders are an essential part of the interactive process of conversation and explains why videoconferencing software, both commercial (e.g. See-it, ) and non-commercial [11], focuses on providing a face-to-face perspective.

For over thirty years, videoconferencing has continued to be a marketing success waiting to happen. Why has it not been successful? Perhaps its purpose should be redefined: specific tasks rather than general purpose functions as it is now commonly used for. This paper reviews videoconferencing literature to draw lessons from its faults. I will try to answer the question: Why is video not advantageous for group collaboration, despite the availability of visual information?

The history of videoconferencing began with the AT&T Picturephone that was publicly introduced at the 1964 World Fair and was predicted to be the replacement for the voice telephone by the 1970s [3]. AT&T had previously used videoconferencing for its annual stockholder’s meetings and found the results successful [7]. Molnar [9], a vice-president at Bell Laboratories found visualising non-verbal cues such as a smile or a non-response while using the Picturephone beneficial. However, it was a marketing flop. One of the reasons cited for its failure was that users felt self-conscious about being a TV character [3].

The general consensus is that the first phase of videoconferencing technology was a dismal failure. Although expected to reduce travel and energy consumption, it was not perceived by consumers to be a substitute for long-distance travel nor was it seen as an energy saving alternative during the OPEC crisis of the ‘70s [3]. Perhaps people prefer human contact to collaborating with video (socialisation over isolation).

Early failure of videoconferencing can be attributed to two areas: the needs assessment methodology and the unyielding desire to replicate face-to-face meetings without fully understanding the function of video [3, 10, 13]. Attitude surveys which ask participants to judge a device that they have never seen or used can produce questionable results. [3]. For example, people may want a talking car (e.g. „your door is ajar“, „your lights are on“), but
when such cars were introduced in the 80’s they were marketing failures. Video can only provide similar features from face-to-face conversations, it cannot replace it as was originally intended [3, 10].

The current phase has gained more acceptance. Sellen [13] confirms this, citing as proof the number of video media centres and video systems which exist or are being created in research centres around the world. Edigo [3] states that success with videoconferencing is found only when used as a supplement when face-to-face communication is not practical, not as its replacement. He outlines specific examples of successful use of videoconferencing. They are: national sales staff training, using frequent meetings in order to make deadlines and the inclusion of junior members in meetings with size restrictions due to budget limitations. These successes may give more weight to the opinion of Edigo[3] that what determines success or failure of video is its application rather than the type of videocommunication system. However, his predictions of increased videoconferencing due to technological shifts towards the office and increased technological competence of employees have not held up.

Although there are perceived advantages to video, I come back to the original question: Why is video not advantageous for group collaboration? For the answer, one must examine factors that contribute to positive group collaboration. These are: field of perception, sense of presence, shared frame of reference and movement. Within each of this areas, the use of video disrupts the opportunity for effective group collaboration.

Field of perception is one’s perceptual field of view that enables one to see objects in one’s periphery that are not directly being attended to. Peripheral cues are used during interpersonal communication to recognise changes in body or head movements and then respond appropriately [6]. A video camera cannot provide a peripheral perspective as only a small part of the environment is visible limiting awareness.

A video camera’s location can either restrict detail when providing a wide room perspective (i.e. top corner of room) or limit perspective (i.e. above computer screen) while focusing on fine details such as a face. In addition, the video screen resolution, limits document size, restrict ability for close-ups, and causes one to miss subtle communication cues (i.e. glances) and more noticeable ones (i.e. hand waving) if one is not observant [6].

Sense of presence, highest when physically present, is the feeling of being part of a group. Sellen [13] after examining 3 multi-user videoconferencing systems including picture-in-picture (most common), found that participants didn’t feel a sense of presence. This led to problems of floor control, that was reflected, for example, when speakers named the person that should next speak, an explicit method of turn-taking. This behaviour she attributes to individual perception that gaze and gestures necessary for floor control are not as effective when interacting with video [see Heath and Luff, [5] for empirical support or Isaac & Tang [6] for similar results].

A shared frame of reference occurs when individuals share the same physical space. Many cues and conversational information about other individuals are provided by a shared reference, such as: Who am I looking at?, Who is looking at me? The areas examined are turn-taking, side conversations, movement and pointing gestures.

In Sellen [13], participants reported that in conversations, it is important to be able to attend individually to others and know when others are attending to them. Gaze and gesture are used in face-to-face conversations to indicate whom one is speaking to and to signal whom the next
speaker should be [12]. They may also be used to prevent another from gaining the floor [2]. However, as observed by Isaacs [6] and also Sellen [13], it can appear as though one is looking at everyone when directly looking into the camera rather than a specific participant. One can also see a problem when group members use words such as ‘you’.

Videoconferencing reduces or eliminates one’s ability to regulate turn-taking, as gestures or other visual cues may be difficult to see, unavailable or perceived ineffective by participants. Isaacs and Tang [6] found that participants did not exchange turns as frequently as when face-to-face which they attribute to participants not being able to tightly regulate utterances and in turn reach a common understanding. Side conversations, (i.e. whispering with neighbours to clarify discussion points), are impossible because the only audio channel is shared eliminating the ability to speak to a particular individual rather than the general group [6, 13].

Movement is an important part of conversation (i.e. pointing out documents, physically changing locations and gesturing to another participant). Gaver [4] considers movements essential for workplace collaboration and perception. With a video screen, three dimension information is lost because the two dimension video eliminates stereopsis, convergence (only one image available) and movement parallax (changes in environment produced by independent movement of observer). Therefore, information necessary for discriminating objects and differentiating the movement of people is limited [4].

Pointing is problematic whether at a document or at a screen because of the difficulty in determining where someone is pointing. Exploratory movement is not possible with stationary cameras, because participants have only one angle to view a shared work environment so physical movements towards different group members are not possible [4].

Although video provides visual information it is a problematic alternative to face-to-face communication because it is restricted by design features such as image size, location of video camera and the audio availability. These design concerns create problems with signals such as gaze and gestures essential for regulating the flow of conversation in group collaboration but easy to miss when not attending to that particular area on the screen. What is needed is an alternative approach that removes that obstacles associated with video while maintaining the sense of ‘being there’ currently present in face-to-face conversations.

References