GColl: Enhancing Trust in Flexible Group-to-Group Videoconferencing

Petr Slovák
Masaryk University
Botanická 68a
Brno, Czech Republic
slovak@ics.muni.cz

Pavel Troubil
Masaryk University
Botanická 68a
Brno, Czech Republic
pavel@ics.muni.cz

Petr Holub
Masaryk University
Botanická 68a
Brno, Czech Republic
hopet@ics.muni.cz

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
In this paper we describe a quantitative study of a group-to-group videoconferencing environment called GColl that provides a compromise between the need for preserving non-verbal cues and the requirements of low-cost and flexibility. We have compared the task process and outcome of participants interacting over an environment analogous to common commodity solutions, those using face-to-face communication, and groups communicating over GColl. Our results demonstrate that it is possible to design a group-to-group collaboration environment with modest technical requirements and low overall cost that still shows measurable advantages over the common environment in its ability to support trust in social dilemma games.

Keywords
Video Conferencing, Gaze Awareness, Evaluation, GColl

ACM Classification Keywords
H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces - Computer-Supported Cooperative Work

General Terms
Design, Performance, Measurement

Introduction

In the last few years, videoconferencing environments and tools came once again into the focus of researchers studying human-computer interactions. This is evidenced by an increasing number of papers on the topic inside the academic sphere itself (e.g., [2,4,7]), but also from the success of telepresence and other less sophisticated videoconferencing systems in the commercial world.

Several recently published papers have also successfully challenged the idea of videoconferencing systems having none, or only mediocre, advantages over audio (e.g., [3]), and showed that interaction over disparate videoconferencing environments might achieve markedly diverse results both in task outcomes as well as task process [4] – thus suggesting that different videoconferencing designs might heavily influence the interaction among the users.

While many systems have been successfully proposed to support videoconferencing among single-participant endpoints, little work has focused specifically on the problem of *group-to-group* videoconferencing, where the preservation of gaze awareness and other non-verbal cues become even more difficult, requiring the creation of specialized systems.

An example of a videoconferencing environment that specifically supports these interactions is Multiview [4]. Even though the Multiview design seems to be very effective (it provides for a "full spatial faithfulness", and the authors have reported positive both self-report, and task-based results), the environment is not well suited for teams that require the possibility of frequent changes in the number of attendants, do not have a special room to place the structure of cameras and the viewing screen in, or are created only for a short-term tasks, thus meeting just a few times altogether.

To support these types of "*ad-hoc*" groups, we have created a videoconferencing environment, called GColl, in which we have tried to reach a compromise between the need for preserving non-verbal cues, and the requirements of low-cost and flexibility. We previously described GColl [6], and argued that it supports eye-contact (based on the same principles as in Multiview) and partial gaze awareness for all participants, while keeping only modest technical requirements. Additionally, users can join or leave GColl on the fly without any changes to the physical layout on any site.

In this paper we present the initial results of an experimental evaluation of GColl based on a moderately large study (90 participants): in particular, we focus on the differences in trust relations appearing among users interacting over GColl, face-to-face, or an environment analogous to the design of current commodity videoconferencing systems (i.e., one screen, one camera per group).

Our results demonstrate that it is possible to design a group-to-group videoconferencing environment with modest technical requirements and low overall cost that still shows measurable advantages over the common environment in its ability to support trust in social dilemma games.

Related Work

Trust creation is considered to be exceedingly important for successful collaboration in group settings [1] and also seems to be heavily influenced by the used communication environment (as was shown in, e.g., [1,4,5]). Additionally, it seems that trust behavior can be measured by observing process and outcomes of social dilemma games [1,4] – and it is, therefore, well suited for experimental research. Trust is in these cases usually defined as, e.g., "[willingly] increasing one's vulnerability to [the actions of] another, whose behavior is not under control" [5].

For instance, Nguyen and Canny (2007) reported a difference in task outcome between Multiview, a videoconferencing environment with supplied gaze awareness information, and the "standard" environment [4]. Based on their results, interaction over Multiview is, in some ways, statistically indistinguishable from face-to-face communication. According to the authors, this is due to the added gaze awareness information in Multiview, which is missing in the "standard" videoconferencing environment. In a study conducted by Bos et. al. (2002), the task outcome was statistically non-significant when face-to-face, video, and audio conferencing were compared; but there were great differences in the task process [1]. Similar results, usually even stronger for other environments such as chat, pure audio or mail interaction, are shown also by other studies (see, e.g., [5,7,8]).

In contrast to the Multiview study [4] where the groups at each site played as single units (thus, the trust was measured between the two remote groups), we focus on changes in trust for each individual separately, while still keeping the group-to-group form of interaction. The used task itself might be of interest also for other studies.

Evaluation Study

Overview of the Experiment

The purpose of the reported experiment is the comparison of three disparate communication media: the *standard videoconferencing* environment (SVE), the *GColl* environment, and *face-to-face*.

Groups of six participants communicated over a particular medium while using a social dilemma game as the basis for their interaction; each group used only one medium during the experiment. In the computer-mediated settings, the participants were divided into two 3-person groups communicating over the videoconferencing environment; in the face-to-face setting, all participants sat around one oval table. We have used the information from game logs as indicators of intra-group trust.

Based on the previous studies on trust in computer-mediated settings, which seem to suggest that the inclusion of non-verbal cues (especially gaze awareness information) leads to higher levels of trust, we expect that: (a) trust differs among groups using different media; (b) trust is highest in groups interacting face-to-face; and (c) trust is higher in groups using GColl than in groups using SVE.

Subjects

Overall, 90 participants (i.e., 15 groups of 6) took part in the experimental part of this study, all were students of various Czech universities. We have taken a great care not to include more than 2 people who have known each other before in each of the groups (in the majority of the groups, all were strangers to each other). We have also inquired about previous use of other videoconferencing software (no participant has

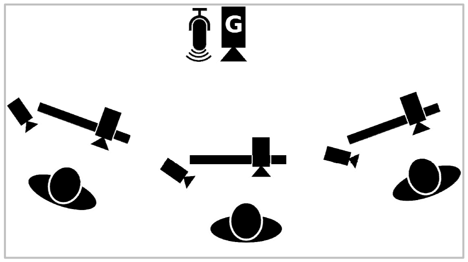

Figure 1: GColl site configuration

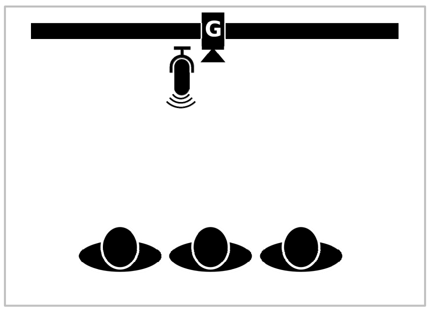

Figure 2: “standard” site configuration

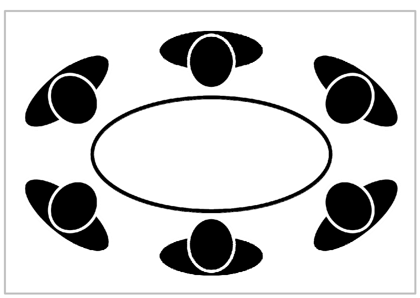

Figure 3: face-to-face configuration

indicated anything more than “using Skype once in a while”).

Media Conditions

Each GColl user was equipped with a common notebook with integrated widescreen display or a PC equipped with a common LCD. Video streams from the users were captured by two Logitech QuickCam Pro 9000 USB webcams at each workstation. At both sites, the whole group view was captured by an Elmo PTC-15S camera. Sound was recorded by a ClearOne AccuMic PC microphone at one site and SHURE EasyFlex EZB/C microphone at the other (see Figure 1 for diagram of the configuration at each site; and [6] for a detailed description of the GColl design).

In the “standard” videoconferencing environment, whole room camera and the sound were captured analogically to GColl, and a NEC MT1060 projector was used to visualize the image of the remote group. The screen was exactly in front of the sitting group in the distance of approximately 3m from the participants in the height corresponding to a normal sitting position; the visualized images of remote participants were slightly smaller than life-size (see Figure 2).

In the face-to-face condition, all participants sat around an oval table, each using a laptop with the game application (as shown at Figure 3).

Game Description

Goldminers game, used in the experiment, is an enhanced version of the well-known Prisoner’s Dilemma task. In this game, participants represent goldminers and try to mine gold from a river. The differences between the Goldminers game and a classical multiplayer Prisoner’s Dilemma include:
(1) communication among the participants is allowed between rounds; (2) there are three possible actions to be taken instead of two; (3) the meaning of these actions differs depending on the context (e.g., an *illegal mining* action may be an uncooperative action if everyone else plays *legal mining*, but may be a cooperative action if the whole group decides to mine *illegally*); and (4) the payback for individual actions changes (in a known way) during the game.

While these differences make the game progress slightly harder to analyze in some aspects, we believe that it demands greater amount of coordination among the players should they try to achieve the cooperative scenario (see below), as a quite complex strategy needs to be created early in the game if the group is to succeed. Additionally, we believe that this version of the game should be, therefore, closer to real-world situations.

In the game, the river has an attribute called the *gold density* which is set to $30,000 at the beginning of the game. At the beginning of each round, each participant chooses one of three possible actions: (a) *legal mining*, which gives the participant lower personal profit (current value of *gold density* minus 25 % tax) and causes no harm to the others; (b) *controlling* the river costs the participant a small amount of gold ($15,000 divided evenly among all participants controlling that round) while incurring a great loss to all illegal miners; (c) *illegal mining*, which is either worth $50,000, if there was no one controlling in the current round, or causes illegal miner to lose the same amount of money in case a control action was chosen. In both cases, *gold density* attribute is decreased by a $1,000 for each

	GColl	SVE	FTF
end-game	294	142	370
variances	15.42	20.84	3.17
controls	7.40	10.20	0.40

Table 1: Mean values of trust indicators (end-game scores in US$ thousands)

	GColl	SVE	FTF
end-game	82.80	12.86	96.72
variances	10.88	8.72	7.35
controls	4.56	3.83	0.55

Table 2: Std. deviations of trust indicators

illegal miner. Once every three rounds the river partially "cleans itself" by increasing the attribute by a $1,000. After all participants choose their action, the round is evaluated and the numbers of particular actions taken (but not who actually took them) are displayed. The game ends after 15 rounds, or if the *gold density* attribute is ever lower than $1,000. Participants were aware of the exact game ending conditions.

Two ending scenarios were possible: if at least 5 out of 6 participants had more game money than a given threshold ($330,000), participants were awarded with a chocolate bar each; on the other hand, if at least 2 participants did not have enough gold, the group members were given chocolate bars according to their results (first two participants got two bars, the next two got just one bar, and the last two did not receive any bars). Thus, incentives for cooperative as well as uncooperative play were present.

Note that the threshold value was quite high (it was just slightly lower than the score gained by 15 rounds of choosing nothing else than legal mining actions) so the participants had to cooperate quite extensively if they were to reach it at the end of the game.

Indicators of Trust

Several game based variables might be used as indicators of trust. A quite crude (but a widely used) one are the *end-game scores of individual participants* (employed in, e.g., [4]). The rationale behind this is that, as a group, the best average score can be achieved only if the group cooperates; and a relation of mutual trust is needed for any cooperative act as, at all times, an uncooperative action yields a large individual bonus to the defector.

Another indicator of trust relations in the game process is the *variability* of group scores for each round. Turn variability tells us how close the turn scores of individual participants in the group are for that particular turn: if the group cooperates, everyone should choose similar actions (implied by the game rules) resulting in low variability. If, on the other hand, someone defects (e.g., playing a *control* when the group decided to play *illegal mining* actions) the resulting variability is high for that turn.

The third possible indicator is the number of *controls* – the only personal gain this action might give is the possibility of a relative increase of one's position (due to others losing even more); also, it can be played to discourage others from defecting, indicating distrust.

Results

To analyze the difference among means for these indicators, we have performed one-way nested ANOVA using the medium of communication as the fixed factor, and the groups as a nested random factor. A normal (i.e., non-nested) ANOVA is inappropriate as the interaction inside each group might have a strong effect on the outcomes on its members. Means and standard deviances for all three indicators are reported in Table 1 and Table 2, respectively. ANOVA tests were significant for all indicators, $p<.05$, (see Table 3 for the F-values), which allowed us to test planned repeated contrasts (i.e., comparing face-to-face vs. GColl; and GColl vs. "standard" environment). All these contrasts were also significant and the effect values (r^2) for these tests are reported in Table 4.

	F-value
end-game	F(2;12) = 7.29
variances	F(2;12.452) = 2.82
controls	F(2;12) = 10.63

Table 3: F-values for the ANOVA tests

	FtF-GColl	GColl-SVE
end-game	0.16	0.44
variances	0.18	0.29
controls	0.72	0.56

Table 4: Effect values for the planned contrasts

Discussion and Future Work

All three operationalizations of trust that were used in this experiment – i.e., end-game score (which can be understood as task outcome), turn variances and the number of control actions (representing the task process) – have supported the three proposed hypotheses, especially those expecting GColl achieve better performance than the "standard" environment.

The reported results are therefore quite optimistic as to the effectiveness of interaction supported by GColl, suggesting that even with limited resources, a videoconferencing environment might operate significantly better than the "standard" environment, i.e., a system analogical in design to current commodity videoconferencing products (such as, e.g., TANDBERG Profile series and many others).

We believe that this is mainly due to the enhanced abilities of providing personal intimacy cues (such as the eye-contact information etc.), which is the main difference between the two compared videoconferencing environments.

There are several parts of the conducted experiment in progress: in addition to administering questionnaires (measuring social presence, and usability), we have been also interested in qualitative analyses of the game interactions over individual media (based on the videotaped recordings), and a smaller sized (36 participants) follow-up study pursuing these questions further is currently underway. These results will be reported in a subsequent paper.

Additionally, we are currently working on a parallel problem of extending the GColl design by a deixis support functionality that would employ the advantages of having co-located groups at each site.

Acknowledgements

This project has been supported by a research intent "Optical Network of National Research and Its New Applications" (MŠM 6383917201) and "Parallel and Distributed Systems" (MŠM 0021622419). We would like to thank Erik C. Hofer for his helpful suggestions and support.

References

[1] N. Bos, J. Olson, G. Olson, and Z. Wright. Effects of four computer-mediated communication channels on trust development. In *Proc. CHI 2002*, pp. 135-140.

[2] J. de Vasconcelos Filho, K. Inkpen, and M. Czerwinski. Image, appearance and vanity in the use of media spaces and video conference systems. In *GROUP09*, pp. 253-262.

[3] A. Kramer, L. Oh, and S. Fussel. Using linguistic features to measure presence in computer-mediated communication. In *Proc. CHI 2006*. pp. 913-916.

[4] T. G. Nguyen and J. Canny. Multiview: improving trust in group video conferencing through special faithfulness. In *Proc. CHI 2007.* pp. 1465-1474.

[5] E. Rocco. Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact. In *Proc. CHI98.* pp. 496-502.

[6] P. Slovák, P. Troubil, and P. Holub. Gcoll: A Flexible Videoconferencing Environment for Group-to-Group Interaction. In *INTERACT*'09. pp. 165-168.

[7] Q. Zhang, W. Qu, and K. Zhang. Do strangers trust in video-mediated communication? In *Proc. IWIC'09*. pp. 325-328.

[8] J. Zheng, N. Bos, J. Olson, and G. Olson. Trust with out touch: jump start trust with social chat. In *Proc. CHI 2001.* pp. 293-294.