# MultiView: Improving Trust in Group Video Conferencing Through Spatial Faithfulness

**David Nguyen**
nguyendt@eecs.berkeley.edu

**John Canny**
jfc@cs.berkeley.edu

Berkeley Institute of Design
University of California, Berkeley
Berkeley, CA 94720-1774

## ABSTRACT

Video conferencing is still considered a poor alternative to face-to-face meetings. In the business setting, where these systems are most prevalent, the misuse of video conferencing systems can have detrimental results, especially in high-stakes communications. Prior work suggests that spatial distortions of nonverbal cues, particularly gaze and deixis, negatively impact many aspects of effective communication in dyadic communications. However, video conferencing systems are often used for group-to-group meetings where spatial distortions are exacerbated. Meanwhile, its effects on the group dynamic are not well understood. In this study, we examine the effects that spatial distortions of nonverbal cues have on *inter-group* trust formation. We conducted a large (169 participant) study of group conferencing under various conditions. We found that the use of systems that introduce spatial distortions negatively affect trust formation patterns. On the other hand, these effects are essentially eliminated by using a spatially faithful video conferencing system.

### Author Keywords

CSCW, CMC, Trust, Social Dilemmas, Video Conferencing, Spatial Faithfulness, Gaze Awareness, Eye Contact

### ACM Classification Keywords

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces — Computer-Supported Cooperative Work; H.4.3 [Information Systems Applications]: Communications Applications — Computer Conferencing, Teleconferencing, and Videoconferencing.

## INTRODUCTION

The goal of any computer-mediated communication system is to enable people to effectively accomplish the task at hand. Video conferencing systems, in particular, have often been touted as the replacement to face-to-face communications.


*CHI 2007*, April 28–May 3, 2007, San Jose, California, USA.


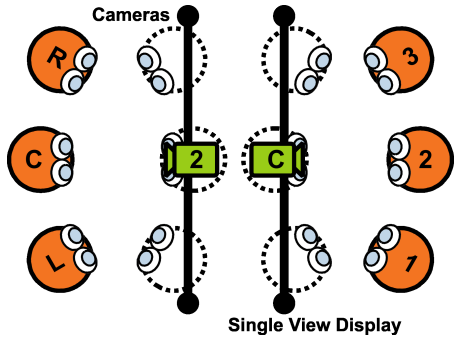


**Figure 1. A typical video conferencing setup for group-to-group meetings. A single camera and single viewpoint display are generally used. Dashed lines indicate images of the remote participants.**

However, many video conferencing systems do a poor job of preserving nonverbal cues that are important in group activities [3, 5, 10, 11, 15, 16] possibly undermining the group's overall goal. In our work, we aim to understand how video conferencing systems affect various aspects of group-to-group communication to inform both appropriate use of video conferencing and design of future systems. In this work, we consider the effect of video conferencing systems on trust formation.

Many nonverbal cues, including gaze and deictic gestures, are dependent on the *spatial faithfulness* of the video system, by which we mean the extent to which it preserves spatial relationships. We previously described the design of a spatially faithful system for group-to-group meetings called MultiView [10], showing that users of MultiView can correctly identify gaze and gesture direction across the video boundary. In this follow-up work, we show that spatial faithfulness is decisive in influencing inter-group trust formation by comparing MultiView, conventional video, and face-to-face meetings in a moderately large study (169 participants). In pursuing this question, we had to adapt previous studies of inter-user trust to the group-to-group case. Our formulation of the trust study may be of independent interest.

Video conferencing systems are notorious for the spatial distortions they introduce. Consider the group-to-group meeting in Figure 1. The dotted characters represent the images of the remote participants on the screen. Suppose

Participant L gazes at Participant 2 on the screen whose image appears right under Camera 2. In this case, Participant 2 will register correctly that Participant L is looking at her. However, Participant 3, who also observes participant L from the perspective of Camera 2, will also perceive Participant L gazing directly at him. In fact, all participants on that side - Participants 1, 2, and 3 - will take on the shared perspective of Camera 2 independent of the actual viewing angle and thus each participant will simultaneously register direct eye contact with Participant L.

We note that spatial distortions and consequent loss of gaze reciprocity apply when a conventional shared display is used *at any site* in a conference. Although we conducted our study on the group-to-group case, we expect similar differences to occur in group-to-individual or group-to-multiple-individuals (at different sites) conferences as well.

It has been shown that video conferencing systems can reduce levels of trust [3]. We extends those findings and presents two results. First, we show that using a conventional shared-screen video conferencing system also reduces trust compared to face-to-face meetings in group-to-group settings. And second, we show that using a spatially faithful video conferencing system – like MultiView – does not produce a significant loss of trust compared to face-to-face.

## PRIOR WORK

### Trust

Many studies have shown that it can be difficult to develop trust using a wide variety of computer mediated communication systems. For instance, Drolet and Morris show that dyads playing a conflict game tended to show more cooperative behaviors when communicating face-to-face than when communicating over the telephone [7]. Rocco showed that 6-person groups playing an investment game tended to show more stable and cooperative investments when communicating face-to-face than when communicating over non-anonymous mailing lists [14]. Bos et al. had 3-person groups play an investment game across four different communication channels: face-to-face, video-conferencing (including audio), audio-conferencing, and instant messenger. They found that participants communicating face-to-face showed higher and more consistent levels of cooperative investment than those using computer-mediated communication systems [3].

In all these previous studies, the structure of the experiment was single participant sites. However, for video conferencing, it is common to have multiple participants at any given site. Our study looks at the effects of video conferencing systems on group-to-group communications for which we have found limited precedence.

### Video Conferencing Systems

Hydra [15] supports multi-party conferencing by providing a camera/display surrogate that occupies the space that would otherwise be occupied by a single remote participant. Because each person has his own individual camera and each remote participant is represented by his own screen, it is possible to direct gaze or gesture at a particular participant and have the recipient register it correctly. GAZE-2 [16] was designed to support gaze awareness for multi-user video conferencing. GAZE-2 uses an eye tracking system that selects, from an array of cameras, the one the participant is looking directly at to capture a frontal facial view. The selected view is presented to the remote user that the participant is looking at, so that these two experience realistic eye contact. Views for the other participants are synthesized by rotating the planar frontal views of the other participants to simulate looking away. MAJIC [11] produces a parallax-free image by placing cameras behind the image of the eyes using a semi-transparent screen.

All three systems above are designed to support multiple single-participant sites. Both Hydra and GAZE-2 support three or more participants, but each participant requires his own setup. MultiView was designed to support multiple sites with individuals or groups at each site.

Nakazawa et al. explored an arrangement of lenticular optics and half silvered mirrors to create “Private Displays” in order to support group video conferencing [9].

## MULTIVIEW DESIGN

In face-to-face, group-to-group communication, each participant in the meeting has his own unique perspective defined by his position. However, video conferencing systems usually only have a single camera whose output video is shared by all remote participants with a single view display. No matter what angle the remote participants view the display from, they will all take on a shared and incorrect perspective defined by the position of the camera. This is known as *perspective invariance* and its cognitive mechanisms are well understood [17]. A related consequence of perspective invariance is the *Mona Lisa Effect*, celebrating the eerie effect of Mona Lisa’s eyes following you as you walk around. This problem will always be present wherever there are multiple participants looking at a shared single-view display.

In order to solve this problem, MultiView adopts a multiple viewpoint directional display that can simultaneously display different video streams to different participants based on their viewing position. Multiple cameras are used to capture unique perspectives for each participant. If the cameras are arranged so that the geometry of a face-to-face meeting is preserved, then each person will see a unique and correct perspective providing full spatial faithfulness for all participants in the meeting. The MultiView arrangement is illustrated in Figure 2. By virtue of the MultiView display, when Participant 3 looks at the display, she sees the video captured by Camera 3. Simultaneously, Participant 2 will see the video captured by Camera 2 and Participant 1, Camera 1. So when Participant L gazes at Participant 2, Participant 2, viewing through Camera 2, sees direct eye contact, Participant 3, through Camera 3, sees Participant

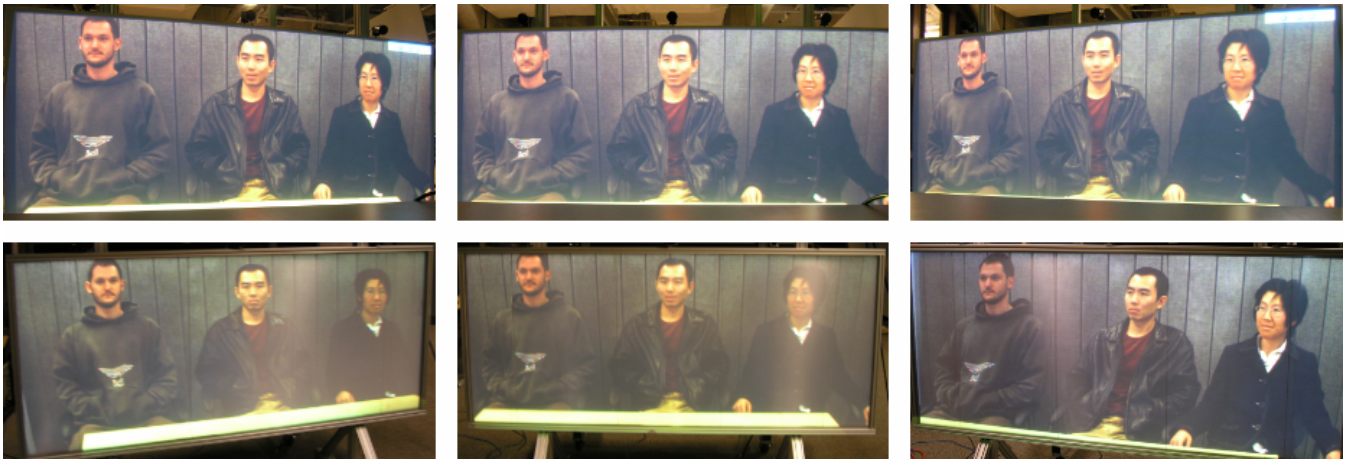

**Figure 3. Three remote participants are gazing at viewing position 1 (see Figure 2). Column 1 is the view from position 1, column 2, position 2, and column 3, position 3. The top row is what is seen from the respective positions with non-directional video conferencing and demonstrates perspective invariance. The bottom row is what is seen using MultiView and shows appropriately changing perspectives.**

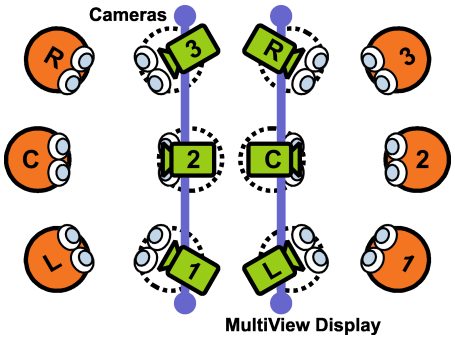


**Figure 2. A MultiView video conferencing setup for group-to-group meetings. In this setup, three cameras are used to capture three unique perspectives which correspond to the unique and correct perspectives of the remote participants. A multiple viewpoint, directional display is used to allow each remote participant to view their respective perspectives simultaneously. Dashed lines indicate images of the remote participants.**

L gazing to his left, and Participant 1, through Camera 1, gazing to the right.

Figure 3 compares what each of the three local participants would see between MultiView and non-directional video conferencing when the three remote participants gaze toward Participant 1. Column 1 is the view from position 1, column 2, position 2, and column 3, position 3. The top row is what is seen using a non-directional video conferencing system like that of Figure 1. The bottom row is the view seen if using MultiView. As can be seen, the top row demonstrates perspective invariance; that is, from all viewing positions, it appears that the remote participants are looking one position to the left. Using MultiView, the perspective changes for each viewing position to correctly show the remote participants looking at position 1.

MultiView went through several design iterations as shown in Figure 4. The first iteration (upper left) explored multiple viewpoint directional displays. It used a modification of the video-tunnel technique [5] combined with a retroreflective fabric screen. Most people using this display found the tunnel too "confining" for video conferencing use and the video quality to be low. Additionally, it would be difficult to capture images of the participants sitting behind the video tunnels.

The second iteration (upper right) introduced new display optics to avoid having to use a video tunnel while improving image quality. It is also the first fully functioning prototype of MultiView as a video conferencing system. We have presented the design and evaluation of this system in our previous work [10]. Several findings were presented, but the ones relevant to improving the design of MultiView are as follows:

- the setup of the system separated groups by 18' which is much further than face-to-face groups would meet.
- image quality was still not good enough for perception of precise eye contact.
- projectors placed on the table produced a lot of noise and heat, making it uncomfortable for prolonged meetings and blocking the valuable work surface.

From the above findings, we developed the latest iteration of MultiView (bottom). MultiView currently supports three participants per site. Participants sit in front of a conference table about 8' from the screen. Each viewing position is separated by 27" or 16° with respect to the screen.

MultiView now features a bigger, wider screen (72"W x 32"H, 9:4 aspect ratio ) so that we can use life-sized images. Though the basic optical functionality is the same as in our prior work [10], we use higher precision optics in this iteration which greatly enhances the image quality. To complement the new screen, new short-throw XGA (1024x768 pixel) projectors allow us to reduce the viewing distance from 18' to 8'. The new projectors are mounted above the participants, clearing the work surface

**Figure 4. This figure shows the several iterations of MultiView's design. MultiView started as a video tunnel setup with a retroreflective screen (upper left). The video tunnel was abandoned for a front projected directional screen (upper right). The latest iteration includes a larger screen, higher precision optics, short throw projectors, and high resolutions cameras.**

and directing heat and noise away from the participants. To capture the images, new high resolution (1024x768 pixel) firewire cameras replace CCTV cameras. Due to the mismatch between the screen's 9:4 aspect ratio and the 4:3 aspect ratio of the projectors and cameras, the image is much higher than necessary. As a result, we discard the lower 40% of the pixels in both the cameras and projector. Sound is recorded using a single echo cancelling desktop conferencing microphone. Speakers are mounted on the top of the screen. All audio and video are encoded and decoded using MPEG-2 codecs (constant bit rate of 6Mbps/video stream) and transported over a local gigabit network – although the bandwidth used is less than 40Mbps.

Cameras are placed to minimize the vertical disparity between the cameras and the images of the eyes. The difference is generally 6" above. Given that the participants are viewing the screen from about 8', there will be about a 3.6° disparity between the actual gaze direction and the perceived gaze direction in the downward direction. However, even with this disparity, people should still register correct eye contact given that it is below the angular threshold beyond which people perceive a break in eye contact (about 5° in the downward direction) [6].

## DAYTRADER: A COOPERATIVE INVESTMENT TASK

The trust measurement task in our study is an instantiation of a social dilemma game called Daytrader and was originally developed by Bos. et. al [3]. Social dilemmas put participants in a situation where the payoff is higher for defecting than for cooperating with the other participants, but where everyone is better off if everyone cooperates than if everyone defects. Social dilemma games have been used as a measure of trust in a variety of experimental studies [1–3, 7, 8, 14]. Even though there are several well known shortcomings with experiments based on social dilemmas, most involving the generalizability of its findings [13], results are well understood and readily comparable to a large body of prior work.

In this study, we used a modified version of Daytrader to measure levels of trust in group-to-group communication. The rules of the game are as follows:

- There are 2 groups, each group consisting of 2 or 3 participants.
- The groups will play an unknown number of rounds.
- In each round each group is given 60 credits. Each group must decide how many of their credits to cooperatively invest with the other group (cooperate) and how many they wish to save for themselves (defect).
- For each round, a new group leader should make the final decision as to what the investment is going to be.
- The cooperative investment is put into a fluctuating market which will average 50% return over the course of the entire game.
- The earnings from the cooperative investment are divided evenly among the two groups, regardless of each groups' contribution to the cooperative investment. Each group is told how much they earned, but they are not told what the other group earned.
- After every 5 rounds a "Rich Get Richer" bonus is awarded to the two groups. 60 credits are placed into the fluctuating market. The earnings are divided between the two groups such that the proportion of the awarded bonuses is equal to the proportion of the groups' earning in the previous 5 rounds.
- Discussion is allowed at any point in time, either with groupmates or with the opposing group. However, groups have about one minute between each round. After the bonuses are awarded at the end of 5 rounds, the groups are given extra time and are encouraged to have a discussion. Groups are not allowed to share precise numerical investment and earning amounts with the other group.

This game differs from the one presented by Bos et al. [3] by using a fluctuating market which adds noise to the information available to the groups. The goal was to make it ambiguous as to whether returns were the result of the other group's action or the market performance. A fluctuating

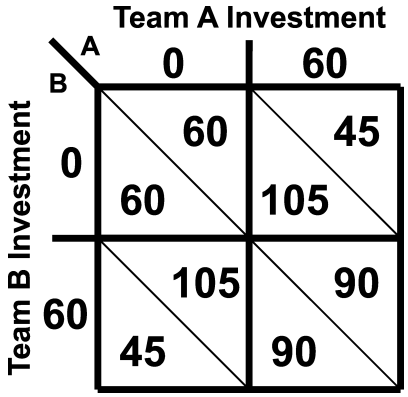


**Figure 5. The payoff structure of Daytrader in four scenarios assuming average market earnings.**

market provides a way to hide defection moves as well as sabotage cooperative moves. It was an attempt to induce more dependence on the communication channel. The participants are made aware that the market is guaranteed to earn 50% on top of the investment by the end of the game and encouraged not to invest based on what they think the market is going to do but on what they think the other group is going to do. The fluctuation was determined before the experiment and was the same across all sessions. The market was determined using a random number generator with an even distribution between -50% to 150% averaging 50%. By adding noise, the game structure becomes an instantiation of what is known as Iterated Prisoner's Dilemma with Imperfect Monitoring [2].

Secondly, the original formulation of this game called for each participant to make his own decision about their investments. In our formulation, groups need to decide how much to invest. To enhance group behavior while reducing effects of dominant and freeloading behaviors, we required that each round has a new group leader who was in charge of making the final decision.

Though a group can decide to invest any amount between 0 and 60 credits, we illustrate the game with 4 possible scenarios (Figure 5) assuming average market performance. If both groups invest 0 credits (upper left), each group will earn 60 credits for that round since they both just saved their credits. If both groups invest all 60 credits cooperatively (lower right), both groups will earn 90 credits. If Group A invests 60 credits while Group B makes a defection move and invests 0 (upper right), then Group A earns only 45 credits while Group B earns 105 credits and vice versa (lower left).

As can be seen from these examples, by investing, a group puts itself at risk for defection by the other group resulting in less earnings than if they invested nothing at all. Additionally, by defecting, they also have the chance to earn more if the other group decides to invest cooperatively. The rational choice is to consistently defect. But once both groups settle on this strategy, both groups will earn less than if they invested irrationally – hence the dilemma.

## HYPOTHESES

In this experiment, we specifically make the following hypotheses based on previous findings:

*Hypothesis 1 (H1): Groups meeting face-to-face will demonstrate higher levels of trust than groups meeting through non-directional video conferencing systems.*

There is limited precedence in measuring trust formation in video conferencing conditions for the group-to-group structure. Finding support showing a difference between face-to-face and non-directional video conferencing conditions in group-to-group meetings adds credence to the problem we are trying to solve. It also provides a basis for comparison. Specifically, we hypothesize the following:

*Hypothesis 1a (H1a): Groups meeting face-to-face will show higher levels of* overall *trust than groups meeting through non-directional video conferencing.*

That is, we expect that the total cooperative investment by groups meeting face-to-face will be significantly higher than the total cooperative investment by groups meeting through non-directional video conferencing.

*Hypothesis 1b (H1b): Groups meeting face-to-face will show reduced* delay *in trust formation when compared to groups meeting through non-directional video conferencing.*

Bos et al. found that trust generally increases over time. They called this phenomenon *delayed trust* [3]. They found that trust increased more slowly with participants meeting through video conferencing compared to groups meeting face-to-face. We expect to extend their results to the group-to-group setting and show that there will be a greater delay in trust formation for groups meeting through non-directional video conferencing when compared to groups meeting face-to-face.

*Hypothesis 1c (H1c): Groups meeting face-to-face will show reduced* fragility *in trust formation when compared to groups meeting through non-directional video conferencing.*

Bos et al. also found that there was a decrease in cooperative investment when bonuses are about to be offered. They called this phenomenon *fragile trust* [3]. They found that trust in participants meeting through video conferencing was less resilient to bonuses than for participants meeting face-to-face. We expect to extend their results to the group-to-group setting and show that groups meeting through non-directional video conferencing will exhibit more fragile trust than groups meeting face-to-face.

Our second hypothesis compares the trust formation patterns of groups meeting through directional vs. non-directional video conferencing systems. We expect that full spatial faithfulness provided by MultiView should improve trust by preserving many of the nonverbal cues which are distorted in non-directional video conferencing systems. Similar to Hypothesis 1, we make the following hypotheses:

*Hypothesis 2 (H2): Groups meeting through directional video conferencing will show higher levels of trust than groups meeting through non-directional video conferencing.*

Specifically, we hypothesize the following:

*Hypothesis 2a (H2a): Groups meeting through directional video conferencing will show higher levels of* overall *trust than groups meeting through non-directional video conferencing.*

*Hypothesis 2b (H2b): Groups meeting through directional video conferencing will show reduced* delay *in trust formation when compared to groups meeting through non-directional video conferencing.*

*Hypothesis 2c (H2c): Groups meeting through directional video conferencing will show reduced* fragility *in trust formation when compared to groups meeting through non-directional video conferencing.*

## METHOD

### Participants

Participants were recruited by the Experimental Social Science Laboratory (XLab) at University of California, Berkeley. The XLab maintains a database of university affiliated students and staff members who are interested in taking part in experiments. Participants are emailed about experiments and opt-in by signing up via an online calendar. There were 169 participants: 110 females (65%), 59 males (35%), 156 students (92%), and 13 staff members (8%). The average age of student participants was 20 years old, and the average age of staff member participants was 39. These participants formed 29 groups of 2 and 37 groups of 3. Groups played against each other in three different conditions in a between-group study.

The experiment occurred in two hour sessions with between four to six participants. Because not everyone always shows up to their scheduled sessions, up to ten participants were recruited for each session. If we could not accommodate a participant in a given session, they would be compensated with a $5 show-up fee. Participants taking part in the experiment were paid according to the outcome of the experiment, but were guaranteed at least $22.50.

### Treatment Conditions

*Face-to-Face:* In this condition, the two groups met in the same room. One group sat on one side of the conference table and the other group sat on the other side. The two groups were separated by 8'.

*Directional Video Conferencing*: In this condition, the two groups met in separate rooms and communicated through the MultiView video conferencing system which takes advantage of the multiple viewpoint directional display and represents a spatially faithful video conferencing system. The groups sat 8' from the screen to mimic the distance of the face-to-face condition.

*Non-Directional Video Conferencing*: This condition was identical to the directional video conferencing condition except the multiple viewpoint display was covered with a standard projection screen material and only the center camera and projector was used. Image quality remained the same. This condition mimicked the commonly found, spatially distorted video conferencing system.

### Measurement Instruments

#### *Task Performance Measure*

The two groups played the variant of Daytrader as described above. The measure of trust is operationalized as the sum total of cooperative investments between the two groups. In each round, the measure of trust can be from 0 (both groups invest nothing) to 120 (both groups invest all their credits).

#### *Post-Questionnaire*

An adaptation of Butler's Conditions of Trust Inventory [4] was administered to the participants. The original inventory consisted of 110 Likert scale questions measuring 11 different conditions. Questions were selected and modified from this pool for appropriateness of the condition to be measured and brevity of the questionnaire. The conditions chosen were *trust in other* group (11 items), *trustworthiness* (5 items), and *consistency* (3 items). This inventory included questions like "I trusted the other group members in this game," "I could be trusted by the other group," and "During the game I behaved in a consistent manner." The participants responded on a scale of 1 (strongly disagree) to 7 (strongly agree).

#### *Post Interview*

Upon completion of the experiment, we interviewed each group separately. There were no predetermined questions, but the topics covered were general impressions of the other group, any specific incidents in the game that stood out (e.g. where they lied, where they felt lied to, etc), and discussion of any strategies they used. The post interview was to help explain some observed events during the game and to guide future research.

### Procedure

The experiment took 120 minutes for each session. Upon arrival, each participant was immediately assigned to one of two groups. If participants were acquainted with another participant, they were placed in the same group. In the computer mediated conditions, participants were escorted to their assigned rooms to minimize any face-to-face contact with opposing group members.

Once assigned to their groups, they were shown a set of videos which walked them through the consent materials and the rules of Daytrader. This process took about 30 minutes.

If applicable, the video conferencing systems were turned on and connected at this point. The participants were allowed to introduce each other to the other group and were given time for discussion before the game began. Once they were ready, they would submit their investment amounts to a *fund manager*.

The fund manager is a program designed to prompt the groups for their investments. The groups interacted with the fund manager through America Online's (AOL) Instant Messenger (IM) program installed on a laptop on their conference table. Once the fund manager received the amounts, it would calculate each group's earnings and report them to the respective groups. Groups did not know the opposing group's earnings. The researcher could command the fund manager to send a "time out" warning, indicating to the participants that they are taking too long to make their decisions. This was necessary to get through enough rounds in the allotted experiment time.

This portion of the experiment lasted for 45 minutes. All groups played at least 30 rounds. The actual number of rounds played was variable between each session and groups were not made aware of how many rounds there would be.

Once the end of the game was reached, the two groups were allowed to say goodbye to each other. In the video conferencing conditions, the systems were shut down and connections were severed. In the face-to-face condition, the groups were separated into different rooms.

Each participant then filled out the post-questionnaire individually and a post interview was conducted. This took about 30 minutes.

The participants were compensated for their participation in the study. The amount of their compensation was based on the number of credits their group earned and the number of rounds they played. Basing the compensation on the number of credits earned during the session provided motivation to do well in the game. The average compensation was $26.21, the maximum was $31.42, and the minimum was the guaranteed $22.50 even if the credits their group earned was worth less.

Each group left at separate times as to avoid meeting again.

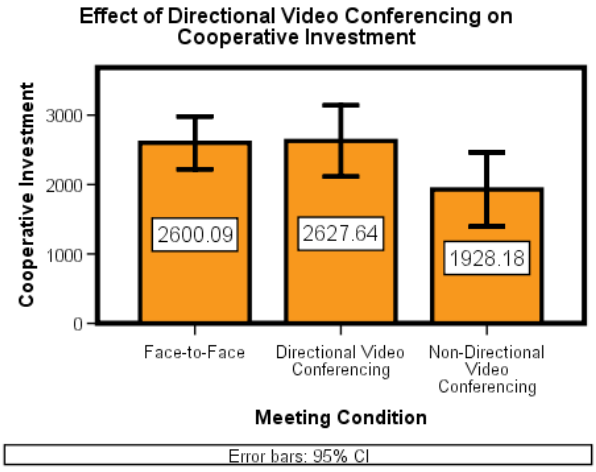


**Figure 6. Overall cooperative investment by meeting condition.**

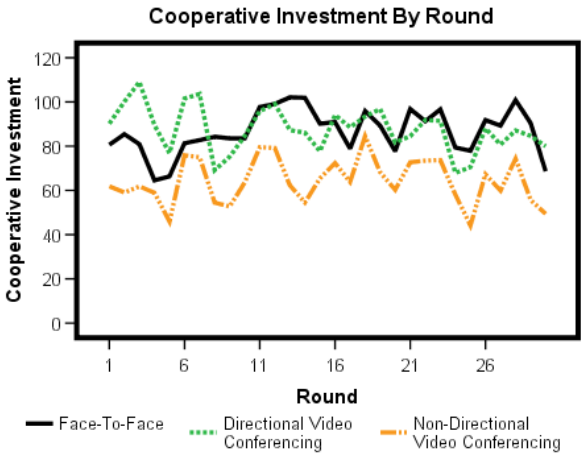


**Figure 7. Cooperative investment amounts by round.**

## RESULTS AND ANALYSIS

### Overall Cooperative Investment

We begin by looking at *overall trust* which is measured by the total cooperative investment across the entire game. We sum all cooperative investments by both groups for the first 30 rounds of each session. The maximum cooperative investment is 3600 credits (60 credits/group * 2 groups/round * 30 rounds).

Means for cooperative investment are shown in Figure 6 for each of the conditions. We performed three Planned Comparisons using one-way analysis of variance. The analysis showed *cooperative investment* by groups meeting face-to-face was significantly higher than by groups meeting through non-directional video conferencing, $F(1, 20) = 5.21, p < .05$. It also showed that cooperative investment by groups meeting through directional video conferencing was significantly higher than by groups meeting through non-directional video conferencing, $F(1, 20) = 4.42, p < .05$. No significant difference in cooperative investment was found between groups meeting face-to-face and groups meeting through directional video conferencing, $F(1, 20) = .01, p = .92$.

### Round-By-Round Cooperative Investment

Next we take a look at investments made round-by-round. For each round, we sum both groups' cooperative investments. The maximum cooperative investment per round is 120 credits (60 credits/group * 2 groups). Figure 7 shows the average of all cooperative investment for rounds 1 through 30. Each line represents a different meeting condition.

Prior work [3] and our data presented in Figure 7 suggest Daytrader data exhibits two different phenomena: (1) *delayed trust*, which is a function of the number of rounds since the start of the game, and (2) *fragile trust*, which is a function of the number of rounds since the last discussion.

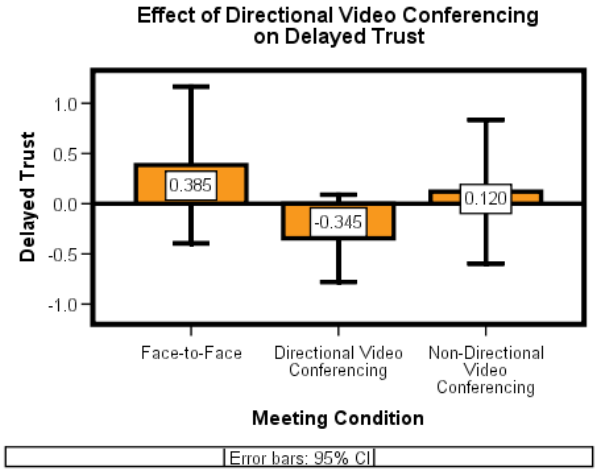


**Figure 8. Delayed trust by meeting condition.**

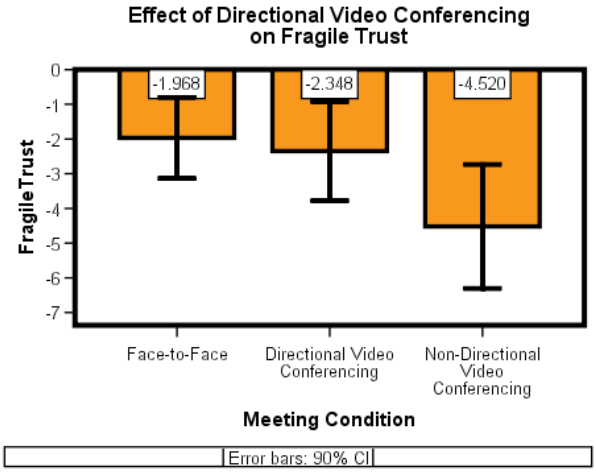


**Figure 9. Fragile trust by meeting condition.**

To use fragile trust in statistical analysis, Bos et al. [3] defined a new variable which we also use called *discussion distance*. It is the number of rounds since the last 5-round discussion. For example, round 6 and 11 both occur right after a discussion, so both rounds would have a discussion distance of 1.

The measure of *delayed trust* is the slope of a regression line between *cooperative investment* vs. *round*, and we calculated it for each of the 33 sessions played. Discussion distance was added as a covariate to control for the effect of fragile trust. The means are presented in Figure 8. We performed three Planned Comparisons using one-way analysis of variance. The analysis showed no significant difference in delayed trust for any of our comparisons: (1) face-to-face vs. non-directional video conferencing, $F(1, 20) = .31, p = .58$, (2) directional vs. non-directional video conferencing, $F(1, 20) = 1.53, p = .23$, and (3) face-to-face vs. directional video conferencing, $F(1, 20) = 3.31, p = .08$. One-way t-tests show none of the delayed trust measurements were significantly different from 0: (1) face-to-face, $t(10) = 1.10, p = .30$, (2) directional video conferencing, $t(10) = -1.77, p = .11$, and (3) non-directional video conferencing, $t(10) = .37, p = .72$.

The measure of *fragile trust* is the slope of a regression line between *cooperative investment* vs. *discussion distance*, and we calculated it for each of the 33 sessions played. Round number was added as a covariate to control for the effect of delayed trust. The means are presented in Figure 9. We performed three Planned Comparisons using one-way analysis of variance. The analysis showed trust in groups meeting face-to-face were significantly less fragile than in groups meeting though non-directional video conferencing, $F(1, 20) = 4.70, p < .05$. It also showed that trust in groups meeting through directional video conferencing was significantly less fragile than in groups meeting through non-directional video conferencing at a more experimental level, $F(1, 20) = 2.96, p < .10$. No significant difference in fragile trust was found between groups meeting face-to-face and groups meeting through directional video conferencing, $F(1, 20) = .14, p = .71$.

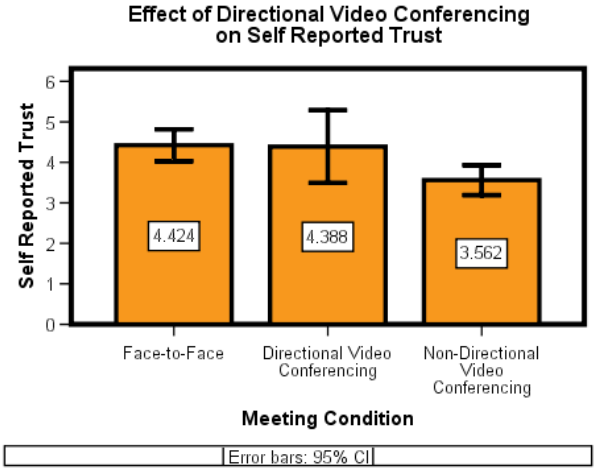


**Figure 10. Self-reported trust by meeting condition.**

### Post-Questionnaire

For each session, the responses to each questionnaire item given by all the participants from both groups were averaged to create an aggregate session response. We disregarded one questionnaire since that participant just circled the same number for all items. The questionnaire measured *trust in other group* (11 items, $\alpha = .96$), *trustworthiness* (5 items, $\alpha = .92$), and *consistency* (3 items, $\alpha = .65$). One item was removed from *consistency* to improve the internal consistency (2 items, $\alpha = .70$). Spearman rank correlation showed a significant and positive correlation between each of the conditions measured in the questionnaire with the total cooperative investment of the game: *trust in other group*, $\rho(31) = .57, p < .01$, *trustworthiness*, $\rho(31) = .50, p < .01$, and *consistency*, $\rho(31) = .48, p < .01$.

We then compared self-reported trust measures by meeting condition. The means are presented in Figure 10. We performed three Planned Comparisons using one-way analysis of variance. The analysis showed that groups meeting face-to-face self-reported significantly higher trust than groups meeting through non-directional video conferencing, $F(1,20) = 12.61, p < .05$. It also showed that groups meeting through directional video conferencing self-reported significantly higher trust than groups meeting through non-directional video conferencing at a more experimental level, $F(1,20) = 3.60, p < .10$. No significant difference in self-reported trust was found between groups meeting face-to-face and groups meeting through directional video conferencing, $F(1,20) = .01, p = .94$. No significant differences were found between experimental conditions for any of the other questionnaire conditions measured. The results of these comparisons match the results of the comparisons of overall and fragile trust from the Daytrader measurements.

## DISCUSSION

On the basis of the above findings, we will now revisit the hypotheses set out earlier.

For this experiment, we used a variant of Daytrader as our measure of trust. The results from the trust inventory show a positive and significant correlation between *investment amounts* and *trust scores* adding internal validity to Daytrader as a trust measurement device.

*Hypothesis 1a (H1a): Groups meeting face-to-face will show higher levels of* overall *trust than groups meeting through non-directional video conferencing.*

This hypothesis is supported on the basis of the descriptive statistics. In comparing the total cooperative investment amount by both groups in all 30 rounds, we found that the total cooperative investment by groups meeting face-to-face was significantly higher than those by groups using non-directional video conferencing.

Additionally, results from the post-questionnaire trust inventory show a statistically significant difference in the *trust in other group* condition between the face-to-face and non-directional video conferencing.

*Hypothesis 1b (H1b): Groups meeting face-to-face will show reduced* delay *in trust formation when compared to groups meeting through non-directional video conferencing.*

This hypothesis is not supported by our results and analysis. Our results do not suggest a difference in delayed trust formation between face-to-face and non-directional meeting conditions. In fact, no change in trust was measured at all in either of these conditions. This may be the result of either a lack of the delayed trust phenomenon in group-to-group interactions or limitations in the power of our experiment for measuring delayed trust. Further studies would be needed to clarify this.

*Hypothesis 1c (H1c): Groups meeting face-to-face will show reduced* fragility *in trust formation when compared to groups meeting through non-directional video conferencing.*

This hypothesis is supported by our results and analysis. Our results show that groups meeting face-to-face tended to be more resilient to breakdowns in trust when compared to groups that met through non-directional video conferencing.

*Hypothesis 2a (H2a): Groups meeting through directional video conferencing will show higher levels of* overall *trust than groups meeting through non-directional video conferencing.*

This hypothesis is supported on the basis of the descriptive statistics. In comparing the total cooperative investment amount by both groups in all 30 rounds across the three meeting conditions, we found that the total cooperative investments by groups using directional video conferencing was significantly higher than those by groups using non-directional video conferencing and that the investments in the directional video condition tended toward the investments made by those who met face-to-face when compared to non-directional video.

Additionally, results from the post-questionnaire trust inventory show a statistically significant difference in the *trust in other group* condition between the directional and non-directional video conferencing conditions. The self-reported trust level in the directional video conferencing condition tended toward the levels in the face-to-face condition which is in agreement with our hypothesis.

From our results, those that met face-to-face invested an average of 2600.09 credits. By using non-directional video conferencing, average cooperative investment reduced to 1928.18 credits, a reduction of 26%. Meeting through directional video conferencing system restored the average cooperative investment up to 2627.63 credits, similar to face-to-face levels. We are careful here not to claim that using a directional video conferencing system like MultiView will fully restore trust lost in using non-directional video conferencing systems, but we do present the lack of measurable difference as support for the dependence of trust on spatial faithfulness.

*Hypothesis 2b (H2b): Groups meeting through directional video conferencing will show reduced* delay *in trust formation when compared to groups meeting through non-directional video conferencing.*

Similar to the discussion for H1b, this hypothesis is not supported by our results and analysis. Our results do not suggest a difference in delayed trust formation between directional and non-directional meeting conditions. No change in trust was measured at all in either of these conditions.

*Hypothesis 2c (H2c): Groups meeting through directional video conferencing will show reduced* fragility *in trust*

*formation when compared to groups meeting through non-directional video conferencing.*

This hypothesis is supported by our results enough to warrant further exploratory work. Our results show a statistically significant difference in fragile trust between the directional and non-directional video conferencing conditions at a reduced level of confidence ($p < .10$). Groups that met through directional video conferencing tended to be more resilient to breakdowns in trust when compared to groups that met through non-directional video conferencing. The measure of fragile trust tended toward that of face-to-face and there was no measurable difference between the face-to-face and directional video conferencing conditions.

## CONCLUSION

Video conferencing systems are notoriously bad at preserving the rich language of nonverbal communications. In our prior work, we introduced the design of a new video conferencing system which preserves many of the nonverbal cues lost in standard video conferencing systems by being spatially faithful. In this study, we examine the effects of spatial faithfulness on trust formation in a cooperative investment task and present two results. First, whereas previous studies focused on single-participant sites, we show that the use of standard video conferencing systems can significantly hinder the trust formation process in multiple-participant sites. The group-to-group configurations in our study cooperatively invested less and trust was more fragile when meeting through non-directional video conferencing than when meeting face-to-face. The second result shows that using a spatially faithful video conferencing, such as MultiView, helps improve the trust formation process. Groups meeting through directional video conferencing cooperated more than groups who met through standard video conferencing systems and were more resilient in their cooperation in the face of temptation. For all our measures of trust, there was no measurable difference in cooperative behavior between groups meeting face-to-face and groups meeting through MultiView.

## ACKNOWLEDGEMENTS

We thank the Experimental Social Science Laboratory (XLab) at the University of California, Berkeley for their help in administering the user study. We thank our collaborators at Hewlett-Packard Laboratories, especially Harlyn Baker and Bruce Culbertson, for their generous support and technical contributions to the development of the MultiView system. We thank Nathan Bos for his help in experimental design. We thank the BiD lab members for their general support and the anonymous reviewers for their feedback.

## REFERENCES

1. Axelrod, R. and Hamilton W.D. The evolution of cooperation. *Science 221, 4489* (1981), 1390–1396.
2. Bendor, J. Uncertainty and the evolution of cooperation. *The Journal of Conflict Resolution 37, 4* (1993), 709–734.
3. Bos, N., Olson, J., Gergle, D., Olson, G., and Wright, Z. Effects of four computer-mediated communications channels on trust development. *Proc. CHI 2002*, ACM Press (2002), 135-140.
4. Butler, J.K., Jr. Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of Management 17, 3* (1991), 643–663.
5. Buxton, B. and Moran, T. EuroPARC's integrated interactive intermedia facility (IIIF): Early experiences. *Proc. of IFIP WG8.4*, ACM Press (1990), 11–34.
6. Chen, M. Leveraging the asymmetric sensitivity of eye contact for videoconference. *Proc. CHI 2002*, ACM Press (2002), 49-56.
7. Drolet, A. and Morris, M. Rapport in conflict resolution: Accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. *Journal of Experimental Social Psychology 36, 1* (2000), 26–50.
8. Hardin, R. *Trust and Trustworthiness*. Russell Sage Foundation Publications, New York, 2004.
9. Nakazawa, K., Shiwa, S., Komatsu, T., and Ichinose, S. Proposal of a new eye contact method for teleconferences. *IEICE Trans. on Communications E76-B, 6*, (1993), 618–625.
10. Nguyen, D. and Canny, J. MultiView: Spatially Faithful Group Video Conferencing. *Proc. CHI 2004*, ACM Press (2004), 512–521.
11. Okada, K., Maeda, F., Ichikawaa, Y., and Matsushita, Y. Multiparty videoconferencing at virtual social distance: MAJIC design. *Proc. CSCW 1994*, ACM Press (1994), 385–393.
12. Olson, G. and Olson, J. Distance matters. *Human-Computer Interaction 15* (2000), 139–178.
13. Riegelsberger, J., Sasse, A.M., and McCarthy, J.D. The researcher's dilemma: Evaluating trust in computer-mediated communication. *International Journal of Human-Computer Studies 58, 6* (2003), 759–781.
14. Rocco, E. Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact, *Proc. CHI 1998*, ACM Press (1998), 496–502.
15. Sellen, A., Buxton, B., and Arnott, J. Using spatial cues to improve videoconferencing. *Proc. CHI 1992*, ACM Press (1992), 651-652.
16. Vertegaal, R., Van der Veer, G. and Vons, H. Effects of gaze on multiparty mediated communication. *Proc. Graphics Interface 2000*, (2000), 95–102.
17. Vishwanath, D., Girshick, A.R., and Banks, M.S. Why pictures look right when viewed from the wrong place. *Nature Neuroscience 8, 10* (2005), 1401–1410.