

# Reassessing the Effect of Videoconferencing Features on Trust in Triadic Negotiations

Siavash Kazemian  
*Dept. of Computer Science*  
*University of Toronto*  
 Toronto, Canada  
 kazemian@cs.toronto.edu

Cosmin Munteanu  
*Dept. of Systems Design Engineering*  
*University of Waterloo*  
 Waterloo, Canada  
 cosmin@taglab.ca

Gerald Penn  
*Dept. of Computer Science*  
*University of Toronto*  
 Toronto, Canada  
 gpenn@cs.toronto.edu

**Abstract**—Globalization, lingering threats of pandemics and ever-increasing travel costs have made videoconferencing a necessary tool in many environments, ranging from corporate meetings with potential clients who may not have met previously, to mediated negotiations between antagonists who must necessarily remain at a safe distance. This raises several questions about how videoconferencing affects the efficiency, fairness and trust that is desired of negotiations. Significant research has been conducted on the effectiveness of video-based collaborative environments, and more recently, on how one specific aspect of trust called credibility develops in videoconference-mediated, collaborative, one-on-one social dilemmas. Much less is known, however, about how typical videoconferencing features affect the full spectrum of trust-connoting attributes that are known within the social science literature, or outside the idealized confines of two-party social dilemmas. We report here the results of a broader experimental investigation that has exhibited significance across other dimensions of trust (dependency and expectancy), both objectively and subjectively, in the presence of more videoconferencing features than just video (chat and screen sharing), and in a realistic three-way negotiation. We make recommendations for designers, administrators, and corporate decision makers with regard to appropriately using these features in videoconferencing systems.

**Index Terms**—videoconferencing; trust; technology-mediated negotiation; chat systems; shared displays.

## I. INTRODUCTION

Videoconferencing is often used in settings where participants do not know each other well, and yet need to engage in sensitive transactions. These transactions invariably require some degree of trust, often involve negotiations as part or all of the transaction, and at times also involve the sharing of documents and presentations, or the collaborative editing of a document, such as a license agreement.

Often such meetings are asymmetrical, for example two participants from a large company, each of them being situated in a different location, engage by videoconference with a third potential client who is located in yet another place. Many videoconferencing systems support text-chat functionality that can allow for back-channel private communications in these cases. But there must still be at least some degree of trust. It has been argued that trust is more difficult to establish through computer-mediated communication (CMC) than through face-to-face communication when it lacks body language and other cues that are used to build trust [1]–[4]. It is widely assumed

that streaming live video of the participants will mitigate this problem to some extent. Does it? What about in relation to other features, such as display-sharing or chatting?

The answer turns out to depend on what we mean by “trust” and which CMC features we intend to use. The two studies presented here evaluate three features of videoconferencing systems — all by now staples of that industry — in terms of their effect on several dimensions of trust. They do so, furthermore, in a more realistic setting than a two-party social dilemma — by now, the staple of HCI studies of trust. *Social dilemmas*, such as the well-known prisoner’s dilemma problem, pit two or more players against each other in a task in which the pursuit of rational self-interest by every player leads to an outcome that is worse than if they pursue some other coordinated strategy that contradicts their self-interests. By contrast, our studies involve a realistic, three-party business negotiation, in which participants had competing interests, the outcomes of the negotiation were not a zero-sum game, and the negotiated settlements (or “deals”) were not necessarily symmetrical. Triadic and higher-order interaction effects have been argued [6] to be essential to understanding the dynamics and formation of social networks. Trust has been shown to be a key potential factor in keeping certain triadic, collaborative problem-solving activities on track in the face of lapses in mutual understanding [7]. This paper attempts to revisit earlier experimentation on trust in CMC in view of these more recent developments.

This paper demonstrates that there are significant differences that emerge in risk tolerance and in other dimensions of trust through variations to affordances that are independent of the presence of video. These other dimensions are significantly impaired by the addition of private chatting functionality and significantly enhanced by screen-sharing functionality. Earlier work on video-enhanced versus audio-only interfaces has found the provable benefits of adding high-quality video or viewrange-expanded video to be less straightforward to quantify. Ours is also among the very few studies to date to have observed significance in an objective measure of trust (specifically, risk tolerance) other than efficiency. These results were observable because of a controlled experimental design that eschews social dilemmas and embraces a fuller conception of trust that is already pervasive within the social sciences.

## II. RELATED WORK ON THE CHOSEN AFFORDANCES

Video itself has already received a fair amount of attention.<sup>1</sup> The two most closely related studies to what we undertake here are those of Bos et al. [8] and Teoh et al. [9], both of which we shall consider in detail presently. The former concerned videoconferencing relative to three alternatives; the latter compared two different videoconferencing views.

In qualitative studies [10], video has been argued to “build trust and relationship,” and even to discourage deception [11]. Where measured, the frame rate of the video seems almost totally irrelevant, on the other hand; even showing the other parties’ faces occasionally in a series of still images is enough to achieve the same effect [12].<sup>2</sup> Another study [14] used a social dilemma game to investigate potentially trust-improving warm-up activities, such as casual introductory interactions by email, visual identification through photos and reading dossiers of personal information. Of these, visual identification was found to be the most significant by far.

As for related work on chatting and screen-sharing, there have been other studies of chat systems apart from Bos et al.’s [8], e.g., [15], but they likewise consider chat in the absence of video and audio, or, in one exceptional case, speech synthesized from text chats [16]. This is mainly a reflection of their age; we consider video and audio now to be a minimum baseline that any software for negotiation would incorporate. In other studies, chat, where it is present at all, often augments at least one of video or audio. One study [17] compared face-to-face communication with text chat in both low-interdependence tasks like brainstorming and high-interdependence tasks like negotiation, and found openness and trust to be beneficial to the highly interdependent task. They also found that a wide “temporal scope” in the relationships of participants (i.e., how long they had known each other) mitigates the difference between face-to-face and text-chat communications. Another [18] conversely observed higher levels of trust emerging from both video-mediated and chat-mediated brainstorming tasks than from similarly mediated negotiation tasks. Studies of shared workspaces (called “shared displays” in the literature, although this does not refer to streaming video broadcasts) are comparatively rare. One study [19] observes that shared displays may create a false sense of shared data validity in certain collaborative work environments.

## III. RELATED WORK ON TRUST

In one respect, our present study continues a thread of HCI research that is perhaps epitomized by Bos et al. [8],

<sup>1</sup>An earlier body of research had attempted to disabuse technologists of the proposition that videoconferencing can ever be a direct replacement for face-to-face meetings [5]. Those studies did not vary the interaction mode settings of a videoconferencing system, however. Instead they looked to shortcomings in both needs assessment methods and the argumentation surrounding promised reductions in business-related travel to draw their conclusions, taking the technology of videoconferencing itself to be a *fait accompli*.

<sup>2</sup>From the standpoint of security and deception prevention, on the other hand, photos of models have been found to impair participants’ abilities to identify trustworthy e-commerce sites [13].

who studied differences in trust as outcomes of a two-person social dilemma resolved through videoconferencing, a three-way phone call or text chatting software, relative to a face-to-face topline. Their subjective, Likert-scaled measures of trust are based on Butler [20], together with correlations of those subjective measures to ultimate aggregated payoff. Statistically significant correlations, as a function of the communication modality, were calculated to group payoff (0.53), self-ratings of trustworthiness (0.69) and self-ratings of consistency (0.61). The trust responses were found to have the same profile of pairwise differences as group payoff. There are important differences in baseline selection relative to our study; the alternatives used by [8] were either a phone call or text chatting, but not even audio with chatting. Even today, audio with chatting is still a very common CMC setup for low-bandwidth communication. Otherwise, this was a very thorough study, but a study based upon only one dimension of trust.

### A. What is Trust?

Our own five-dimensional view of trust is based upon the instrument that Alston [21] selected in a study of trust in technology-driven organizations, called the Organizational Trust Index [22], although for clarity we will use less domain-specific terms for the dimensions: expectancy trust, dependency trust (which, following [23], we take to be synonymous with risk tolerance<sup>3</sup>), credibility trust, empathy and competence.<sup>4</sup>

Butler’s [20] account of trust, upon which the measures of trust in [8] are based, is by contrast entirely affective. This kind of trust might better be called trustworthiness, i.e., a partly emotive, entirely human-centred account of what makes people trustworthy, such as, for example, their status as authority figures in a society. This is not atypical of treatments of trust found in the behavioural sciences literature (e.g., [25]). When studying features of communication-mediating technology, however, the design of the experimental scenario should reflect instead on the trust requirements of the situation or context in which the communication takes place.<sup>5</sup> In addition, as Wierzbicki [23] argues, the technology should, to the extent possible, be rooted in an entirely rational or what he terms “cognitive” view of trust in order to be computationally realizable. CMC technology should not strive to be inert, in other words, but rather to be an intelligent component of the communication channel. But in these negotiations, the communications technology is not cast in the role of a would-be human participant. It silently enables and mediates.

Our own multi-dimensional view has allowed us to calculate similar correlation scores to those of Bos et al. [8], but further-

<sup>3</sup>This view of risk considers tolerance to be a contextually variant resultant state, which is conditioned over time by experience dealing with other actors. An alternative view is adopted by [24] in which certain individuals are inherently more risk-prone than others. They find this alternative to correlate positively with an individual’s construal of self as being defined by his or her relationships.

<sup>4</sup>The OTI’s terms are openness/honesty, identification, reliability, concern for employees, and competence, respectively.

<sup>5</sup>See, as an example, Riegelsberger et al.’s [4] taxonomy.

more to do so for each dimension of trust independently. Very few previous CMC studies take trust to be multi-dimensional (versus being a desirable primitive alongside other positive attributes of communication, such as fairness, openness, etc.). Although it would certainly be an overstatement to claim that there is universal agreement in the social sciences literature on what “trust” means, there has been a growing awareness within management science since Butler’s [20] early work on the subject that trust is in fact multi-dimensional. Alston [21] (p. 30) went so far as to include the ability to measure multiple dimensions of trust as one of four selectional criteria for her survey instrument, alongside validity and ease of completion within less than 10 minutes. One finds this sentiment outside of management science as well. Mitchell and Zigurs alone have found 10 unequivocally distinct definitions of trust in other literature [26]. See also Uslaner [27] on the multidimensional nature of trust resulting from factor analysis of trust surveys. Any CMC research that has claimed to measure primitive “trust” according to any single definition, no matter which definition they select, has in some sense failed to capture what trust really is.

### B. Social Dilemmas

Teoh et al. [9] is a rare exception of a CMC study that does recognize trust as multi-dimensional, but it is still based upon a two-personal social dilemma. They were interested in removing the experiential bias that earlier comparisons to face-to-face communication had not satisfactorily addressed, as well as in evaluating the effect of varying amounts of visual information through a combined video/audio/chat interface across two different types of tasks, creative and negotiating. There found a generally greater amount of trust using a wider camera angle,<sup>6</sup> but which type of task exhibited greater trust depended on the specific dimension of trust measured.

Social dilemma games, which have been used by numerous CHI/CSCW papers on the topic of trust over the last 25 years, e.g., [14]–[16], [29], [31]–[34], have been widely criticized in the social sciences literature as inadequate for the study of trust because, *inter alia*, the only cause for distrust in a classical social dilemma game is the fear of defection. This fear can be equated with the negation of *commitment trust*.

Social dilemma games are convenient because they are abstract, zero-sum games, in which it is comparatively easy to conduct controlled experiments. Nevertheless, they are not representative of a great many real negotiations. In the case of Teoh et al. [9], the choice of a dyadic social dilemma tragically prevented them from measuring objective risk tolerance. The ecological setting of the negotiations we use in our study is very similar to theirs, and among our several objective measures of trust are their measures of time to completion and fairness of payout. As will be seen below, the only one in which we found significance was risk tolerance. Fairness

<sup>6</sup>Other papers have studied the use of multiple cameras and camera angles [29], and the relative merits of listener-controlled versus speaker-controlled video cameras [30].

of payout has elsewhere been observed to improve in videoconferencing negotiations when the tasks are more difficult or have more competing trade-offs, and when less information is exchanged per conversational turn [34].

Indeed, most previous work with social dilemmas has focussed on dyadic (two-person or two-group) social dilemmas,<sup>7</sup> in which there is generally no opportunity to measure any kind of trust apart from *credibility trust* because blame is so readily assignable to the other party.

## IV. RESEARCH QUESTIONS AND HYPOTHESES

The main goal of our research was answering the question of whether videoconferencing features, such as chat and screen sharing, influence the (by definition, measurable) attributes of deal-making and the outcomes of multi-party negotiations. In particular, we are interested in the following attributes and their outcomes: participants’ trust in each other, their tolerance for risk, the perceived and real (i.e., time-elapsing or turn-counting) efficiency of the negotiation, participants’ sense of equity (we shall call it; more precisely, it is distributive fairness [23]:p.16) and transparency. All of these attributes pertain to the aforementioned five dimensions of trust from the OTI [21].

To answer these research questions, we formulated several inter-related hypothesis schemata  $\mathbf{H}(\mathbf{Y})$ , where  $Y$  is a variable ranging over (e)fficiency, (r)isk tolerance, (c)redibility trust and e(x)pectancy trust. In addition, we formulated hypothesis subschemata relating to the chat function  $\mathbf{Hc}(\mathbf{Y})$  and the screen sharing function,  $\mathbf{Hs}(\mathbf{Y})$ :

$\mathbf{H}_e(\mathbf{e})$ : Participants’ **efficiency** is affected by the videoconferencing features available for use. Public and private chat leads to decreased perceived and real efficiency of the meeting negotiation ( $\mathbf{Hc}(\mathbf{e})$ ). Screen sharing ( $\mathbf{Hs}(\mathbf{e})$ ) leads to increased perceived and real efficiency.

$\mathbf{H}_r(\mathbf{r})$ : Participants’ **tolerance for risk** is affected by the videoconferencing features available for use, with public and private chat leading to decreased risk tolerance ( $\mathbf{Hc}(\mathbf{r})$ ) and screen sharing ( $\mathbf{Hs}(\mathbf{r})$ ) leading to increased tolerance.

$\mathbf{H}_c(\mathbf{c})$ : Other participants’ perceived **credibility** is affected by the videoconferencing features available for use, with public and private chat leading to decreased perceived transparency ( $\mathbf{Hc}(\mathbf{c})$ ) and screen sharing ( $\mathbf{Hs}(\mathbf{c})$ ) leading to increased transparency.

$\mathbf{H}_x(\mathbf{x})$ : Perceived **expectancy trust** among participants during negotiations is affected by the videoconferencing features available for use, with public and private chat leading to decreased perceived equity ( $\mathbf{Hc}(\mathbf{x})$ ) and screen sharing ( $\mathbf{Hs}(\mathbf{x})$ ) leading to increased equity.

Our hypothesized positive sentiment towards screen sharing followed from [19] and our expectation that screen sharing would likewise create a sense of transparency during negotiations.

<sup>7</sup>See, however, [15] for a pioneering exception, in which a non-anonymized electronic mailing list was evaluated against a face-to-face baseline in six-person social dilemmas. [35] also observes the effects of group formation on trust in three-person groups.

As for the presence of text chat, Birnholtz et al. [36] proposed that it can provide a sufficient situational awareness for completing collaborative problem-solving tasks without necessarily ensuring a complete understanding of the situation. Lee and Tatar [7] showed that, in those same activities, trust could be a key factor in keeping the task on track in the face of lapses in mutual understanding. [7] also showed that dyadic orientations can emerge within a triad that lead to different criteria of understanding among different pairs of participants. Turning instead to triadic negotiation activities, we hypothesized that dyadic sub-orientations would still exist (although we have not tested this hypothesis directly), and therefore that the availability of text chat in the interface would likely lead to a situational awareness of being excluded from one of those dyads that would be sufficient only to causing a breakdown in trust. Because the situations that we constructed are not zero-sum games, this would not be a foregone conclusion; it would say something very contingent about the dynamics of group orientations, negotiation tasks and trust.

### V. EXPERIMENTAL DESIGN

In order to test our hypotheses, we designed a within-subjects study in which participants were exposed to differing subsets of interaction modes, in an official, group-meeting scenario. Three student participants and one mediator negotiate over the inclusion of various facilities into a new recreation centre that will be built on our university’s campus.

#### A. Participants

The study was conducted using 72 students (32 female and 40 male), grouped into 24 meeting groups. All participants were undergraduate and graduate students from various disciplines at the same university. While the maximum age was 45, the median age was 25, with 80% of the participants being between the ages of 18 and 27.

#### B. Independent Variable

**Video Only (VO)** — a baseline setting available to all modern videoconferencing systems, consisting of two-way voice and video interaction.

**Video+Chat (VC)** — the Video-Only setting augmented with a chat interface for exchanging text messages (and no screen sharing). Text messages can be “broadcast” to all participants or sent privately to a single party. There is no defined limit to the lengths of the text messages, although the 33-character width for the chat column would make it awkward for transmissions of more than a paragraph in length. Figure 1 shows the videoconferencing system with these features enabled.

**Video+Screen (VS)** — the Video-Only setting augmented with H.239 dual-stream screen-sharing (and no text chatting<sup>8</sup>). The screen-sharing feature is capable of annotating content and highlighting text on the screen to bring the attention of

<sup>8</sup>A single setting in which all three were available was not investigated in order to avoid confounds from participant preferences and selection of non-audio/video affordances, due to more than one being available.

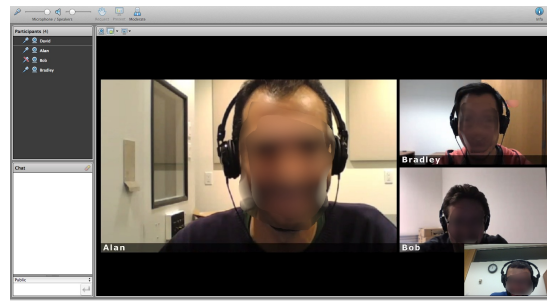


Fig. 1. The videoconferencing system, displaying participants video feeds, the list of participants, and the chat panels (private and public). For privacy considerations, participants’ faces have been blurred.

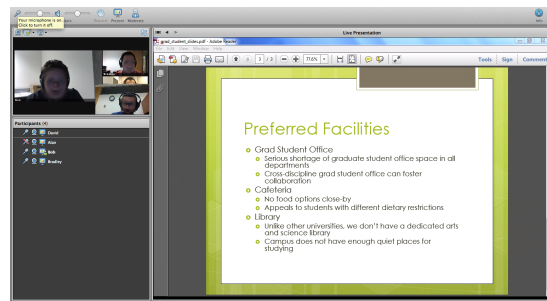


Fig. 2. The videoconferencing system, displaying participants video feeds, the list of participants, and the screen sharing feature. Private and public chat were not functional for this condition.

other participants to it, and also features a “content-slider” that allows viewers to slide back and forth in past time, independently of the presenter’s current state of presentation. Figure 2 illustrates the screen-sharing feature of the videoconferencing application.

All three conditions were rendered using a popular commercial videoconferencing system that runs an IETF Session Initiation Protocol stack (SIP) on a remote server, and broadcasts streaming, high-definition (720p at 30 frames per second) H.264 video, G.722.1 (“Polycom”) audio and other content between browser-based Java clients that each run on their own Windows desktop computer. The overall system requires 1 Gbps channel capacity.

#### C. Task

In both studies, each participant in the experimental trials plays the role of a student representative for one of the large extracurricular societies on campus, who negotiates on behalf of their society for the facilities of a new recreational centre that will be built at the university. Depending on the facilities chosen, there is also the risk of a tuition increase, which subjects must balance against the desirability of the facilities. Given that our subjects are students, this task is ecologically valid because of its direct pertinence to their lives on campus.

#### D. Allocation of Roles and Experimental Conditions

These experiments call for a repeated-measures (within-subjects) design in order to account for the natural variance

that exists in human-subject interaction styles and negotiating abilities. Each experimental session or trial consists of a meeting with four participants. The governing council student representative, an interactor who ran the experiment, was the same in every trial and served as a moderator. The other three participants, the human subjects, participated in their roles as student representatives. The moderator was trained to not intervene in negotiations, and was only active in the initial phase of each trial (describing the task and outlining the scenario, namely the facilities to be discussed and the student groups whose interests were at stake), and as a voting facilitator.

Each group of three human subjects participated in three negotiating sessions with each other (and the moderator), one for each setting of the interaction mode. Every human subject negotiated with the same two other human subjects in his/her three trials, with no other participants added or removed.

A Latin square design of size three was chosen to randomize the order in which participant groups were exposed to the three settings of the independent variable [37]. 8 squares were used in both studies, but not the same squares. The two studies were run independently in order to avoid the exponential increase in time and expense that would have resulted from a single, size-five grid, but the two size-three grids do at least share one common variable setting (VO). The squares were designed so that: (1) each setting of an independent variable was matched with one of three negotiation scenarios an equal number of times, and (2) each of the three scenarios appeared in some position in the sequence given to each group of three participants.

The three scenarios are defined by the three societies represented in their negotiations:

- 1) athletic society, graduate students' society, MBA students' society
- 2) arts society, athletic society, undergraduate students' society
- 3) Christian students' society, health-and-fitness society, undergraduate students' society.

The two societies that appear in more than one scenario (athletic and undergraduates) appear in either instance with different rankings and different tuition-increase judgements.

#### E. Measures and Instruments

For the purpose of comparing the effect of each interaction mode setting, we followed [23] in eliciting a range of both subjective *user perception* judgements and objective *performance* measures that instantiated and integrated our multiple dimensional scales of trust.

These two types of data were collected and measured for each<sup>9</sup> experimental condition in both studies. We used a metric

<sup>9</sup>To maintain ecological validity, the scenarios were not altered in order to "force" participants to use the features that were enabled/disabled for each conditions. As such, participants were free to make use of all features available for each condition: the chat feature was extensively used by all three participants in only 10 of the undiscarded 22 meetings of the video study, while the screen sharing feature was used by all participants in 13.

conjoint analysis [38] on a triadic variant of the actor-partner interdependence model [39], in which each actor evaluates both other partners together. This relies on no individual appearing in more than one group.

1) *User perception data*: We asked participants upon completion of each of the three meeting tasks (one for each experimental condition) to indicate their agreement to statements related to their perception of trust, and the degree of transparency, efficiency, and equity in the meetings. Agreements were represented on a 5-point Likert scale, from 1 for strongly disagree to 5 for strongly agree. The statements included:

- 1) "I could trust the other participants,"
- 2) "The other participants were exchanging important information that I could not see,"
- 3) "The other participants were withholding information from me,"
- 4) "The other participants disclosed all relevant information during the meeting,"
- 5) "I feel that I completed the task in its entirety,"
- 6) "The other participants were effective negotiators,"
- 7) "I had the tools necessary to complete the task efficiently,"
- 8) "I had to work under pressure,"
- 9) "It was difficult to understand the voices of the other participants,"
- 10) "The settlement that we reached was fair to all parties,"
- 11) "I was able to participate actively in the meeting,"
- 12) "The other participants dealt with me fairly,"

as well as a general question related to the interface: "I found the interface intuitive and easy to use." Each statement was rephrased to a statement of the opposite polarity independently with 50% probability to control for bias in the wording.<sup>10</sup> Statements were also presented in a randomly selected order.

2) *Objective performance data*: Our research focus was on measuring participants' attitudes and perceptions with respect to various trust-like attributes. As such, most of the data collected were subjective in nature. However, participants' willingness to engage in transactions that are riskier with respect to the potential for a fairer payout for them is often cited as an indicator of trust [23]. Since, during negotiations, our participants had to advance proposals based on preferences and risks that were assigned to them for each of the facilities (illustrated in Table I), we collected their aggregated tolerance for lower- or higher-risk facility selections. The process for selecting facilities and factoring in the risks is outlined in Section V-F. The value for the **aggregate risk tolerance** of a given session was calculated as the total number of times any

<sup>10</sup>The exception is statement (7), in which the modals "necessary" and "efficiently" were combined in a single statement so that the topic would not be overrepresented in the questionnaire. From among this statement and two rephrasals:

- "I was unable to find all of the information I needed on the interface," and
- "The task required a great deal of effort."

one was independently selected with 33.3% probability.

TABLE I  
AN EXAMPLE RANKED LIST OF FACILITIES, TOGETHER WITH  
ASSESSMENTS OF WHETHER THEY WOULD RISK A TUITION HIKE.

Facility	Rank	Risk of Tuition Hike
Weight Room	1	No
Swimming Pool	2	No
Physiotherapy Centre	3	No
Grad Student office	4	Yes
Cafeteria	5	No
Library	6	Yes
Rentable Office Space	7	Yes
Meeting Rooms	8	No
Virtual Stock Market	9	No

of the four facilities chosen by consensus appeared at risk of a tuition hike in any participant’s facilities table.

In addition to risk tolerance, we instrumented the experiment to collect objective data on the **disparity in the preference rankings** of the facilities ultimately agreed upon by the negotiations, a measure of fairness or equity. Performance-related efficiency measures, such as **time to reach a settlement** and **number of voting rounds** needed to agree on a settlement, were intrinsically available and thus also incorporated into our analysis.

#### F. Balance Conditions

The tables of ranks and risks that were provided to subjects were balanced so that, in a single session: (1) the ranks of each facility were all different on the three tables, (2) the sum of the ranks of any facility was constant, depending on its classification as high importance (sum = 18), medium importance (sum = 15) or low importance (sum = 12; the importance of the facilities was not disclosed to subjects), (3) every society’s top-ranked facility was assigned no risk of a tuition hike in its own table, but a risk of a tuition hike in the tables of the other two subjects in the same scenario, (4) every society’s third-ranked facility was assigned no risk of a tuition hike in its own table, and in the table of one other subject in the same scenario, but a risk of a tuition hike in the table of the third subject, and (5) every society’s second-ranked facility was assigned no risk of a tuition hike in the tables of all three subjects in the same scenario. These balance conditions ensure that some negotiation is required, but that a best solution (consisting of the three second-ranked facilities plus one other) and several near-best alternatives are always available, without resorting to a zero-sum game (as in [9], [10]). Unlike [9], in particular, these balance conditions do not entail that the sum of ranks is always the same, which would preclude our computation of risk tolerance (see below), as the removal of any facility would result in a constant reduction in payout.

This aspect of our design is also important because our coordinated use of facility rank and risk (in our sense of a tuition-hike) actually decouples our framing of risk (in the sense of the dependency trust dimension of Wierzbicki’s scales [23]) into the threat of a realized loss (in the form of a tuition increase) plus the possibility of an unrealized gain (of desirable

facilities being incorporated into the recreational centre), thus addressing, and possibly neutralizing, the bias that either could exert by itself on human risk assessment [40]. It furthermore does so without the use of exogenous auxiliary hypotheses (as in [14]) that may inadvertently convey the impression to participants that their decisions are less consequential or less precise in their effects.

#### G. Rounds of Voting

Successive rounds of voting were conducted until the participants converged on a choice of 4 facilities (each receiving 3 votes), or until 30 minutes had elapsed for the session. Of the 24 groups (=72 sessions) that we convened, 22 reached consensus in all three of their sessions. The data for the 2 meetings without consensus were discarded.

The potential for multiple rounds of voting in the experimental design is also important, as it fundamentally differentiates the structural incentives inherent to this task from those inherent to social dilemmas, which are generally one-shot, highly symmetrical contests, in which both parties must simultaneously act as both potential trustors and potential trustees. The multiple proposals that can be advanced during negotiation, together with potentially multiple rounds of voting makes trust more amenable to analysis as a signalling problem [4]. Ongoing discussions between rounds of voting also create ample opportunities for constructive verbalization about trust and distrust. This has been recommended (again by [4]) as a means of addressing their third dimension, sources of vulnerability. The only vulnerability that players face in social dilemmas is the possibility of defection, which is oversimplistic.

It must be conceded, however, that the potential for multiple rounds of voting was largely underutilized. Of the sessions that ever reached a certain conclusion, 64% terminated after one vote, and 35% terminated after 2, leaving only 1% with more than 2 rounds of voting. On the other hand, there was a substantial amount of verbalization. Of the 129 sessions in the 43 groups across both studies that always reach consensus, 29 (14 from the video study) included pejorative comments relating to trust (mainly complaints of being left out of the discussion, accusations of defection and statements of incredulity), and 36 (24 in the video study) included positive comments relating to trust (mainly affirmations that a recent proposal is fair to everyone, approving statements about the outcomes of previous agreements, and comparative arguments that a participant should agree to a current proposal because they had agreed to a similar one earlier).

#### H. Procedures

Subjects were first screened to confirm their student status and that they were unacquainted with the other two subjects in their group. The two studies were run in non-overlapping sequence (video study first) and no student participated in both studies.

Prior to the commencement of the first session, subjects were indoctrinated into the setting in which all of the fictitious

negotiations took place. Each participant was informed that they would play a different character in each of the three sessions. Each character has a name, and the participants were required to address each other in each session by their character’s name in that session. Participants were also informed that the university would act upon the consensus that they established, and that their re-election as student society representatives would depend crucially on the outcomes of the negotiation. This creates realistic (but fictional) social risks on top of those that pertain to the negotiation’s direct outcomes.

Participants were seated in separate offices on the same floor of an academic building, each of which was equipped with office furniture and a desktop computer running the videoconferencing system. Between sessions, participants took on the roles of new characters from new student societies, but did not switch office spaces.

In each session, the moderator allowed discussion to take place for about 10 minutes, provided the other participants with some positive reinforcement for their efforts and suggested that they vote after another 5 minutes of discussion. After each round of voting, the results of the voting were announced, and, in case the votes were overdispersed, the moderator would select the 5 or more facility choices that received the highest numbers of votes, and encourage more discussion about these, excluding the other facilities from consideration. The moderator never cast a ballot. Note that while the top-ranking facilities were announced, the ballots of individual participants were not, and there were always three voting participants. Again, without the use of exogenous hypotheses, the presence of a third voter allows us to avoid certain attribution of loss in almost every case. Certainly attributable loss is another common shortcoming cited of two-player social dilemma games, as this is often not the case in real negotiations either. At the same time, the presence of a third participant allows pairs of participants to “gang up” in settings where private chatting is available.

At the conclusion of the final ballot in each session, a written post-meeting questionnaire was distributed for each participant to answer by themselves. At the conclusion of the third session’s post-meeting questionnaire, an additional post-experimental questionnaire was distributed to gather additional comparative user perception data and demographic information, followed by an experimental debriefing.

### I. Data Analysis

In order to test our hypotheses, the most suitable statistical test for this within-subjects design is a repeated-measures multilinear model (MLM), a generalization of ANOVA [41] common to quantitative research on the actor-partner independence model of trust [39], that includes an explicit variable for group-level effects, as there almost certainly will be interactions between individual subjects in the same negotiations.<sup>11</sup>

<sup>11</sup>Because no two groups shared an individual, we also speculatively calculated an ANOVA by modelling each group as an individual with subjective preference scores equal to the average of its three individuals’ scores. The results of that analysis were not substantially different from the MLM analysis presented here.

The exception is the objective performance data, which can only be calculated at the group level. For these, we used a simple repeated-measures ANOVA as implemented by the PAST system [42]. F-tests in the MLM were approximated with the Kenward-Roger method using the *pbkrtest* library in the R programming language [43]. All tests were evaluated assuming a size of  $\alpha = 0.05$  for the null hypothesis’s rejection region, and all post-hoc Tukey tests were calculated using a Bonferroni correction quotient. No other transformations were applied to the data. We tested the data for homoscedasticity, but also computed a non-parametric test, Wilcoxon’s signed-rank test, and  $\chi^2$  scores in order to confirm the validity of the F-scores obtained. We also calculated simple descriptive statistics for each setting of the independent variables.

## VI. RESULTS

**Synopsis:** Most of the hypotheses **Hc(Y)** and **Hs(Y)** are confirmed by the experimental results. The analysis of both the objective and subjective data collected from the experiment suggests that chat, particularly private chat, does not facilitate a positive environment for conducting negotiations in an equitable, transparent, and risk-free manner (hypotheses Hc(Y)). However, some of these negative effects are reversed when participants use screen sharing as their main collaborative features. Screen sharing also contributes a somewhat positive boost to the efficiency of the negotiations (hypotheses Hs(Y)).

All the subjective data were collected as participant agreement or disagreement with various statements on each post-treatment questionnaire as previously detailed in Section V-E. Our goal was to analyze how each experimental treatment condition (that is, chat or screen sharing) influenced the deal-making attributes of the whole meeting vs. the baseline treatment (video only), yet the subjective data was collected individually for each participant.

### A. Efficiency

Several post-experiment questions elicited participants’ impressions of how efficiently the negotiations proceeded. The MLM on some of the efficiency-related statements shows a significant relation between the videoconferencing features and participants’ perceptions of efficiency ( $F(2,130)=3.77$ ,  $p=0.03$ , and averages illustrated in Figure 3(a) for the “I was able to find all of the information I needed on the interface” statement), confirming hypothesis H(e). Post-hoc Tukey tests showed significant differences both between the screen sharing condition (VS) and the video only treatment (VO;  $p=0.01$ ) and between the chat condition (VC) and VO ( $p=0.03$ ), thus confirming both sub-schemata Hs(e) and Hc(e). The assessment, “I had the tools necessary to complete the task efficiently,” was marginally significant across treatments ( $F(2,130)=2.56$ ,  $p=0.08$ ) but the other subjective assessments of efficiency (“I feel that I completed the task in its entirety,” “The other participants were effective negotiators”) did not show significant differences. Similarly, objective measurements of efficiency, such as meeting duration (viz. time to completion), did not exhibit statistical significance across treatments.



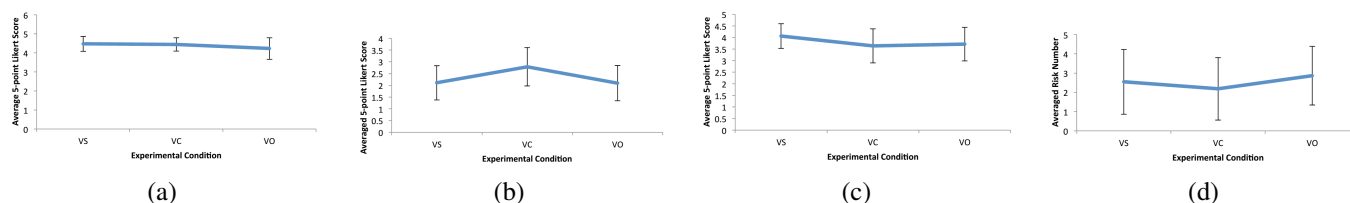


Fig. 3. Responses to the statements: (a) “I was able to find all of the information I needed on the interface,” (b) “The other participants were exchanging important information that I could not see,” and (c) “I felt that the settlement we reached was fair to all parties,” averaged for each of the three experimental conditions: VS (Video+Screen sharing), VC (Video+Chat), and the baseline VO (Video only), along with (d) the risk tolerance value. The variance of the risk tolerance is large, but correct. The effect size was calculated as small (0.030), but significant above Cohen’s threshold of 0.01.

### B. Risk Tolerance

A main objective measure used in the study was the participants’ tolerance for risk. Each participant negotiated under a certain set of preferences for facilities that were to be selected, and a set of tuition fee increases. These increases were presented as yes/no “rumours” of increases to avoid a temptation to simply perform addition on precise tuition increments when determining the overall merit of a proposal. To quantify the risk taken by participants during negotiations, we have tallied for each session the number of times any of the four facilities chosen by consensus appeared at risk of a tuition hike in any participant’s facilities table. As indicated in Section V-D, these tables were not identical across participants, both in the preferred rank of facilities and in the rumoured tuition fee increase. Participants were prohibited from divulging their preferred ranks, but they were allowed to reveal the rumoured tuition increase if needed to move the negotiations along. There were several permutations possible in which participants could have settled for facilities that were favourably-ranked (in the top third of their tables), yet these permutations carried different tuition increase risks. As such, the aggregated risk number was an objective measure of participants’ willingness to opt for a riskier selection of preferred facilities in exchange for a settlement that was somewhat optimal for all participants.

The MLM on the aggregated risk numbers indicates a statistically-significant relation between videoconferencing features and participants’ tolerance for risk ( $F(2,42)=3.271$ ,  $p=0.048$ , and averages illustrated in Figure 3(d)<sup>12</sup>), thus confirming hypothesis H(r). The sub-schema Hc(r) was also confirmed — a Tukey post-hoc test shows that using chat led to participants being less willing to take risks than when using only the video ( $p=0.01$ ). Screen sharing did not exhibit any influence on participants’ risk tolerance as compared to the baseline feature.

### C. Credibility Trust

Several statements on the post-treatment questionnaires were related to transparency or obscurity of the negotiations. Of these, the analysis for the answers to the most salient one

<sup>12</sup>The substantially greater variance in this figure is due to the group-level calculation of objective scores, because of which there is a smaller sample size.

(“The other participants were exchanging important information that I could not see”) indicate a statistically-significant strong influence of the videoconferencing features on participants’ perceived transparency of negotiations ( $F(2,130)=9.73$ ,  $p < 0.01$ , averages in Figure 3(b)), thus confirming H(c). Pairwise Tukey post-hoc tests reveal that chat had a negative influence on transparency compared to both screen sharing and basic features ( $p < 0.01$  for both).

The statement “The other participants were withholding information from me” also exhibited a significant influence by the chat condition on participants’ perceptions of transparency ( $F(2,130)=4.06$ ,  $p=0.02$ ), strengthening the confirmation of Hc(c). The statement, “The other participants negotiated with me honestly,” was marginally significant over the different treatments ( $F(2,130)=2.54$ ,  $p=0.08$ ) but other statements that were less-obviously connected to transparency (e.g., “I was able to understand the motivations of the other participants”) did not exhibit any statistically significant variations.

### D. Expectancy Trust

An important aspect of any negotiation is to have all participants engaged equitably and fairly in discussions, and to reach a mutually beneficial compromise. It relates to trust for the same reason that fairness does (see above). In our study we asked participants if they felt that “the settlement that we reached was fair to all parties.” The MLM analysis on the agreement/disagreement answers shows that these were significantly affected by the videoconferencing features, thus confirming hypothesis H(x) ( $F(2,130)=3.43$ ,  $p=0.04$ , averages in Figure 3(c)). Tukey pairwise comparisons show a significant differences ( $p=0.046$ ) between the effect of screen sharing (VS) and that of both chat (VC) and video only (VO). While the separate hypothesis sub-schema Hc(x) was not confirmed independently, the confirmation of H(x) together with the situation of average responses for VO between VS and VC suggests that chat may have a negative effect on the equity of the negotiation while screen sharing may provide the opposite.

The other equity-related statements (such as “I was able to participate actively in the meeting” or “The other participants dealt with me fairly”) did not exhibit any significant results, perhaps attributable to the overall civilized environment and respectful behaviour of all participants (which may not be always reflected in real-life negotiations).

We also measured equity objectively through a max-min difference, computed by summing the ranks (in the given



preference tables) of the facilities chosen by each participant, and then subtracting the smallest of these sums from the largest. It was also not found to have been influenced by the videoconferencing features.

#### E. Empathy and Competence Trust

No subjective elicitations in the dimensions of empathy or competence trust showed statistical significance.

#### F. Stated Trust

We also directly elicited a subjective judgment on trust itself (“I could trust the other participants”), upon which influence by the videoconferencing features was found to be marginally significant ( $F(2,130)=2.73$ ,  $p=0.069$ ), with a post-hoc Tukey comparison pointing to video plus chat (VC) as the outlier ( $p=0.02$ ). Following Bos [8], we calculated correlations using Kendall’s tau between this judgement and the others, finding values ranging between  $-0.43$  and  $+0.58$ , with the highest correlates all indicating expectancy trust, and the lowest (anti-)correlations all belonging to the negative indicators of credibility.

### VII. DISCUSSION AND IMPLICATIONS FOR DESIGN

The analysis of post-condition questionnaire answers indicates that in several cases the hypothesis schemata  $H_c(Y)$  hold true. This suggests that, in general, chat is a feature that leads to an observed decrease in risk tolerance for participants, as well as to a decrease in credibility trust, expectation trust and stated trust. This is most likely attributed to the possibility for private chatting since negotiations could be conducted simultaneously between ad-hoc pairs of participants with private chat. We can thus state that **private chat in videoconferencing systems can be detrimental to participants’ risk tolerance, credibility trust, stated trust and their perception of fairness**. This prompts us to recommend that:

Designers and administrators of videoconferencing systems disable private chat by default for participants engaged in negotiations.

An interesting future direction is the investigation of appropriate ways to enable or disable private chat, and establishing who has control over such decisions.

As outlined in Section VI, the addition of screen sharing led to a relative increase in perceived efficiency, but not to a significant decrease in actual time to completion. It also led to a relative increase in expectancy trust over the video-only baseline, but not to a significant increase in the objective equity of the outcome. While we did not test screen-sharing in the absence of a mug-shot video interface in either study, it is worth underscoring that the observed benefits did accrue on top of whatever might have been conferred by videos of the other participants. We can thus state that **screen sharing improves the subjective experience of efficiency and fairness**. Therefore, we recommend that:

Designers and administrators of videoconferencing systems consider enabling access to screen sharing features for participants in negotiations.

Where does this leave video? Earlier studies that used two-person social dilemmas generally had no opportunity to measure any kind of trust apart from credibility trust. This is in stark contrast to the addition of screen sharing, for which benefits accrued on top of whatever video may have added, and to the elimination of chat functionality, which the presence of video and audio alone was incapable of ameliorating in any of the several affected categories of trust. Video almost certainly does provide other, marginal, non-trust-related advantages, such as improved vocal intelligibility due to modal enhancement, as well as perceived efficiency due to increased engagement. But it is due to the broader conception of trust in our studies’ design that we have been able to observe a wider range of significant benefits and detriments from screen-sharing and chat functionality. The presence of video, it appears, is by no means the end of the conversation about videoconferencing.

### VIII. CONCLUSION AND FUTURE WORK

Many real-life situations involve multi-party negotiations that cannot be reduced to a zero-sum situation or a social dilemma; instead, the range of outcomes can vary from convergence on one of several possible compromises to an outcome that is very favourable for some participants at the expense of others. Increasingly, such negotiations are carried out by videoconference. To this extent, our experimental setup replicated this real-life situation and thus provided ecological validity to our analysis, as the experimental task was designed to be relevant and familiar to our pool of participants (students, discussing tuition fee hikes and facilities). On the other hand, participants were aware of the role-playing aspect of the experiment, and while most were engaged and passionate about selecting facilities, the lack of real-life stakes may have toned down some of the results, such as equity and how participants treated each other fairly.

This paper evaluated the effect that two ubiquitous features of videoconferencing systems (chat and screen sharing) have on trust, relative to video. They were conducted in a business-like setting in which three participants had competing interests in a negotiation, with possible outcomes ranging from significant disparity between participants’ share of the final settlement to a relatively fair, albeit slightly sub-optimal, deal for all participants. This setting is more realistic than the typical two-party social dilemma that is so often used as a setting for studying trust in videoconferencing systems.

At the same time, our findings provide a very relevant insight into triadic group activities, where the research emphasis has generally been on collaborative problem-solving. Future work should include conversational analysis of realistic group negotiation tasks such as the one pursued here, in order to explore the potential links between interface affordances and group dynamics, such as the dyadic sub-orientations found by [7], more thoroughly. A potential limitation of the present study was the presence of the moderator; triadic negotiations in the absence of a moderator may respond differently to changes in the affordances of the conferencing interface.

The analysis of our experimental results uncovered evidence of significantly negative and positive effects that chat and screen sharing have on various aspects of trust that distinguish them from the presence of video alone. These findings provide useful information for designers, administrators, and decision makers about the appropriate setting for the use of videoconferencing features. Future work should extend this analysis to other settings, such as remote, mobile or home-based participation in meetings. This is of great interest to designers of mobile and tablet-based conferencing systems, where it is understood that one or more participants do not have access to a full-featured application.

#### ACKNOWLEDGEMENTS

This research was supported by the Canadian Network Centre of Excellence in Graphics, Animation and New Media (GRAND) and Avaya.

#### REFERENCES

- [1] S. Whittaker and B. O’Conaill, “The role of vision in face-to-face and mediated communication,” in *Videomediated Communication*, 1997.
- [2] N. Döring, *Sozialpsychologie des Internet: Die Bedeutung des Internet für Kommunikationsprozesse, Identitäten, soziale Beziehungen und Gruppen*. Hogrefe, Verlag für Psychologie, 1999.
- [3] A. Mitra, “Trust, authenticity, and discursive power in cyberspace,” *Comm. ACM*, vol. 45, no. 3, pp. 27–29, 2002.
- [4] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy, “The researcher’s dilemma: evaluating trust in computer-mediated communication,” *Intl. J. Human-Computer Studies*, vol. 58, no. 6, pp. 759–781, 2003.
- [5] C. Egidio, “Videoconferencing as a technology to support group work: a review of its failure,” in *Proc. CSCW*, pp. 13–24, ACM, 1988.
- [6] Z. Fang and J. Tang, “Uncovering the formation of triadic closure in social networks,” in *Proc. IJCAI*, pp. 2062–2068, 2015.
- [7] J.-S. Lee and D. Tatar, “Sounds of silence: Exploring contributions to conversations, non-responses and the impact of mediating technologies in triple space,” in *Proc. CSCW*, pp. 1561–1572, ACM, 2014.
- [8] N. Bos, J. Olson, D. Gergle, G. Olson, and Z. Wright, “Effects of four computer-mediated communications channels on trust development,” in *Proc. CHI*, pp. 135–140, ACM, 2002.
- [9] C. Teoh, H. Regenbrecht, and D. O’Hare, “Investigating factors influencing trust in video-mediated communication,” in *Proc. OZCHI*, pp. 312–319, 2010.
- [10] W. Standaert, S. Muylle, and A. Basu, “Assessing the effectiveness of telepresence for business meetings,” in *Proc. 46th Hawaii Intl. Conf. on System Sciences (HICSS)*, pp. 549–558, IEEE, 2013.
- [11] J. T. Hancock, J. Thom-Santelli, and T. Ritchie, “Deception and design: The impact of communication technology on lying behavior,” in *Proc. CHI*, pp. 129–134, ACM, 2004.
- [12] T. Mvumbi, F. Kundaali, Z. Manzi, K. Williams, and H. Suleman, “An online meeting tool for low bandwidth environments,” in *Proc. Annu. Conf. South African Institute for Computer Scientists and Information Technologists*, pp. 226–235, 2012.
- [13] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy, “Shiny happy people building trust?: Photos on e-commerce websites and consumer trust,” in *Proc. CHI*, pp. 121–128, ACM, 2003.
- [14] J. Zheng, E. Veinott, N. Bos, J. S. Olson, and G. M. Olson, “Trust without touch: jumpstarting long-distance trust with initial social activities,” in *Proc. CHI*, pp. 141–146, ACM, 2002.
- [15] E. Rocco, “Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact,” in *Proc. CHI*, pp. 496–502, ACM, 1998.
- [16] C. Jensen, S. D. Farnham, S. M. Drucker, and P. Kollock, “The effect of communication modality on cooperation in online environments,” in *Proc. CHI*, pp. 470–477, ACM, 2000.
- [17] B. J. Alge, C. Wiethoff, and H. J. Klein, “When does the medium matter? Knowledge-building experiences and opportunities in decision-making teams,” *Organizational Behavior and Human Decision Processes*, vol. 91, no. 1, pp. 26–37, 2003.
- [18] X. Sun, Q. Zhang, S. Wiedenbeck, and T. Chintakovid, “Gender differences in trust perception when using IM and video,” in *CHI Extended Abstracts*, pp. 1373–1378, ACM, 2006.
- [19] B. T. Kane, P. J. Toussaint, and S. Luz, “Shared decision making needs a communication record,” in *Proc. CSCW*, pp. 79–90, ACM, 2013.
- [20] J. K. Butler Jr, “Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory,” *J. Management*, vol. 17, no. 3, pp. 643–663, 1991.
- [21] F. Alston, *Culture and Trust in Technology-Driven Organizations*. CRC Press, 2013.
- [22] P. Shockley-Zalabak, K. Ellis, and R. Cesaria, “Measuring organizational trust: trust and distrust across cultures,” in *The Organizational Trust Index*, IABC Research Foundation, 1999.
- [23] A. Wierzbicki, *Trust and Fairness in Open, Distributed Systems*. Springer, 2010.
- [24] M. K. Lee, N. Fruchter, and L. Dabbish, “Making decisions from a distance: The impact of technological mediation on riskiness and dehumanization,” in *Proc. CSCW*, pp. 1576–1589, ACM, 2015.
- [25] J. K. Burgoon and J. L. Hale, “Validation and measurement of the fundamental themes of relational communication,” *Communication Monographs*, vol. 54, pp. 19–41, Mar. 1987.
- [26] A. Mitchell and I. Zigurs, “Trust in virtual teams: Solved or still a mystery?,” *ACM SIGMIS Database: the Data Base for Advances in Information Systems*, vol. 40, no. 3, pp. 61–83, 2009.
- [27] E. M. Uslander, “Measuring Generalized Trust: In Defense of the ‘Standard’ Question,” in [28], pp. 72–82.
- [28] F. Lyon, G. Mollering, and M. Saunders, eds., *Handbook of Research Methods on Trust*. Edward Elgar Publishing, 2011.
- [29] D. T. Nguyen and J. Canny, “Multiview: improving trust in group video conferencing through spatial faithfulness,” in *Proc. CHI*, pp. 1465–1474, ACM, 2007.
- [30] J. MacCormick, “Video chat with multiple cameras,” in *Proc. CSCW*, pp. 195–198, ACM, 2013.
- [31] E. Bradner and G. Mark, “Why distance matters: effects on cooperation, persuasion and deception,” in *Proc. CSCW*, pp. 226–235, ACM, 2002.
- [32] J. P. Davis, S. Farnham, and C. Jensen, “Decreasing online ‘bad’ behavior,” in *CHI Extended Abstracts*, pp. 718–719, ACM, 2002.
- [33] L. E. Scissors, A. J. Gill, K. Geraghty, and D. Gergle, “In CMC we trust: The role of similarity,” in *Proc. CHI*, pp. 527–536, ACM, 2009.
- [34] W. Dong and W.-T. Fu, “One piece at a time: why video-based communication is better for negotiation and conflict resolution,” in *Proc. CSCW*, pp. 167–176, ACM, 2012.
- [35] P. Slovák, P. Novák, P. Troubil, P. Holub, and E. C. Hofer, “Exploring trust in group-to-group video-conferencing,” in *CHI Extended Abstracts*, pp. 1459–1464, ACM, 2011.
- [36] J. P. Birnholtz, T. A. Finholt, D. B. Horn, and S. J. Bae, “Grounding needs: Achieving common ground via lightweight chat in large, distributed, ad-hoc groups,” in *Proc. CHI*, pp. 21–30, 2005.
- [37] R. E. Kirk, *Experimental Design: Procedures for the Behavioral Sciences*. Brooks Publishing, USA, 1995.
- [38] R. Priem and A. Weibel, “Measuring the decision to trust using metric conjoint analysis,” in [28], pp. 212–225.
- [39] D. L. Ferrin, M. C. Bligh, and J. C. Kohles, “The actor-partner interdependence model: a method for studying trust in dyadic relationships,” in [28], pp. 189–198.
- [40] D. Kahneman, P. Slovic, and A. Tversky, *Judgments under Uncertainty: Heuristics and Biases*. Cambridge University Press, 1982.
- [41] D. C. Howell, *Fundamental Statistics for the Behavioral Sciences*. Nelson Education, 2016.
- [42] O. Hammer, D. Harper, and P. Ryan, “Past: Paleontological statistics software package for education and data analysis,” *Palaeontologia Electronica*, vol. 4, pp. 1–9, 2001.
- [43] U. Halekoh and S. Højsgaard, “A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models—the R package pbrtest,” *J. Statistical Software*, vol. 59, no. 9, pp. 1–30, 2014.