
Challenges in Ethical Design of User Studies with Negotiation Scenarios

Siavash Kazemian

Department of Computer
Science
University of Toronto
kazemian@cs.toronto.edu

Cosmin Munteanu

Department of Computer
Science
University of Toronto
Mississauga
cosmin.munteanu@utoronto.ca

Gerald Penn

Department of Computer
Science
University of Toronto
gpenn@cs.toronto.edu

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced in a sans-serif 7 point font.

Every submission will be assigned their own unique DOI string to be included here.

Abstract

In this case study we describe a human subject experiment that was designed to measure the effects of different features of a video conferencing software on trust established between participants that used it to perform a negotiation task. The fact that each participant can act opportunistically and unethically in the experiment and the social effects on the experiment's payoffs present some challenges to the ethical design of this experiment and those similar to it.

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]:
Miscellaneous

Introduction

An important consideration in the design of any study involving human participants in HCI is to adhere to the highest possible standards of ethics as outlined by [1] and [7]. Ethical conduct of research require the proper elicitation of informed consent, just enrolment process, and managing risk/benefits of the research by the researcher [1]. Although ethical issues related to eliciting informed consent and just enrolment process such as dealing with privacy issue when conducting research using social network data (ex. [2]) or enrolment process of inmates have been well-studied (ex. [10]), issues related to managing the risk and benefits of human subject experiments in HCI have largely revolved

around identifiable vulnerable populations (ex. [13]). However, participants outside of identifiable vulnerable populations can become quite vulnerable during an experiment specially when exposed to actions of other participants [9].

The experiment described in this case study is a group experiment in which participants meet virtually over a video conferencing system to negotiate. This study explores the ethical issues that arise when participants are exposed to each other's deceitful, opportunistic, and otherwise unethical behaviour during the negotiation process. The studied experiment's focus was to examine the role of private chat and screen sharing in trust-building (or trust breaking) between users of video conferencing systems who are negotiating.

Given that the participants were free to act opportunistically and deceitfully, it had to be ensured that such actions did not subject other participants to unacceptable risks. Although aspects related to monetary compensation of participants in such scenarios have been considered (ex. [5]), the social effects in such studies have been less examined.

The dilemma here is that the experiment protocol had to enable participants to act unethically if they so desired, all the while ensuring that unethical behaviour did not expose participants to unmitigated risks. The anticipated social and psychological risk were:

- Reputation risk: Those who acted opportunistically and lied at some point should not gain a bad reputation because of their participation in this experiment
- Self assessment of competence risk: Those who were cheated and achieved lower results in the experiment should not feel incompetent in any way.

But as it will be discussed in this case study, there were risks to the participants that were not anticipated in experiment design and had to be dealt with on the fly by the experiment facilitator. Those risks were associated with emotional and aggressive exchanges between participants who had cheated each other.

Research Background

Previous studies have analyzed the effects on trust of video conferencing and other computer-mediated communication (CMC) systems. Video itself has already received a fair amount of attention [15]. In qualitative studies, it has been argued to build trust and relationships [4] and even to discourage deception [8]. Effects of warm-up activities such as casual introductory interactions by email, visual identification through photos, and reading dossiers of personal information have also been measured based on trust exhibited between participants of prisoner dilemma games over CMC. The most closely related work to what we undertake here is that of Bos et al. [3], who studied the differences in trust and the outcome of a 2-person social dilemma resolved through video conferencing, a three-way phone call, or text chatting software. Lacking is a study on the effects on trust of using private chat while in a video-conference and of screen sharing.

Trust is multi-faceted and affects many aspects of activities between parties. Although most research in HCI has focused on *expectancy trust*, which is "a subjective expectation an agent has about another's future behaviour" [11], and measured it in social dilemma games. But in real-world scenarios in the corporate world, other facets of trust such as *competency trust* and *companion trust* play a significant role in determining the efficiency and risk-taking in transactions [16]. We thus decided to assess the effects of our independent variables on trust by measuring their impact on

efficiency and risk-taking in a more realistic scenario.

The human subject experiment described here was designed to test the following hypothesis about the effects of screen sharing and private chat use on trust by measuring objective and subjective metrics on efficiency and risk-tolerance:

- H1 The use of private chat reduces trust between participants because it can easily be used to facilitate untrustworthy behaviour. This reduction in trust would decrease risk-tolerance amongst participants and meeting efficiency.
- H2 The use of screen sharing would allow participants to expose more of their actions to the rest of the group, thus building trust. The increased trust would lead to higher risk-tolerance amongst participants and meeting efficiency.

Participants met through the video conferencing software to take part in the experiment and, due to the experiment's scenario, at times either acted opportunistically or lost trust in the other participants' goodwill. The experimental design challenge was to create a framework in which participants could act opportunistically, breaking the bonds of trust, all the while shielding them from the negative effects of unethical behaviour that can arise when trust is broken amongst individuals [14].

Experiment

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.

Subjects were first screened to confirm their student status and that they were unacquainted with the other two subjects in their group. Prior to the commencement of the meetings, participants were indoctrinated into the setting in which all of the fictitious negotiations took place. Each participant was informed that they would play the role of a large student group representative in a negotiation meeting attended by two other representatives (also study participants) and a governing council student representative who was the study's facilitator and acted as the meeting moderator. The moderator was trained to not intervene in negotiations, and was only active in the initial phase of each trial to describe the task and to outline the scenario, and as a voting facilitator. Each meeting participant has a fictitious name, and the participants were required to address each other in each session by their character's name in that session. The goal of the meeting was to decide which 4 out of 9 possible facilities to include in the university's to-be-constructed new recreational centre. 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.

The study's independent variable was the means of interaction through the video conferencing interface. We assessed its effects across three different settings:

Video Only (VO): A baseline setting available to all modern video conferencing systems, consisting of a 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 sent to all participants or privately to a single party.

Video+Screen (VS): The Video-Only setting augmented with H.239 dual-stream screen-sharing (and no text chatting). The screen-sharing feature also has a 'content-slider' that allows viewers to slide back and forth in time, independently of the state of the presenter's screen.

To examine the effects of the independent variables on how participants built trust during the meetings, each group of participants took part in 3 negotiation meetings as described above, each time with a different interaction mode. Participants were given new roles and characters in each meeting all of which resembled the described scenario and were asked to forget the previously held meetings.

All three interaction modes were rendered using a popular business-grade conference system that enables high-definition video (and audio) of participants, public and private (one-to-one) text chat, and screen sharing.

The representatives had competing interests in that each facility was ranked differently by each representative. The sum of ranks of each facility by the three representatives was one of three constants. This sum represented the overall importance of the facility to the group. But the participants only knew their own preferences and had to rely on the other participants to find out theirs. Moreover, each participant believed that including some facilities could lead to a tuition hike but there was no consensus amongst participants about building which facilities would lead to this undesired outcome. The participants were given two goals: (1) get their top facilities chosen by the group, and (2) ensure there is no tuition hike. After a negotiation meeting, each participant would vote for 4 out of 9 available facilities and the facilities with the most votes (minimum two votes required) would be included in the recreational centre. In-

conclusive voting sessions would be followed by more negotiations and votes until the group reached consensus on 4 facilities or 30 minutes had lapsed. 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 preferences and beliefs about risk of tuition hike were organized so that there are a set of optimal solutions if the group trusts each other and acts cooperatively. Namely, it is possible for each participant to get their second-highest rated facility (in addition to a fourth medium ranked facility) while choosing only one facility that is believed by one representative of having a tuition hike risk. If the participants don't trust one another, or otherwise want to act opportunistically, they can follow a number of strategies, all of which would resemble a zero-sum game (as in [15]). For example, they could go for their top 3 choices and concede the fourth facility to another representative in order to gain their support for one of their top facilities. In this case, the participant would trust their view of tuition hike risk more than the group's, which would lead to a set of facilities that will have two or more facilities with tuition hike risk from the perspective of two representatives. To ensure the validity of the data in exploring the research questions that were about trust, the experimental protocol had to allow participants to act opportunistically and unethically if they chose to. This could expose the participants in the group to two sources of non-monetary risks¹ that had to be mitigated, namely, reputation risk and self-assessment of competence risk as defined earlier in this paper.

The risk to a participant's reputation was already low given the explicit requirement that participants should never have met beforehand. To further protect participants' reputations,

¹ All participants were paid equally despite the experiment's outcome.

we took care to ensure that participants in the same meetings were not from the same department. Moreover, we ensured that our instructions to the participants made it clear that acting ethically or unethically during the experiment were both normal and admissible behaviours, due to the lack of any tangible effect on participants' life outside of the experiment and thus were not reflective of participants character and tendencies outside of this experiment. This fact was reinforced in our instructions and in a review sessions after the training session.

Just as with other experiments, it is critical to ensure that participants realize their performance in the experimental scenario is not indicative of their capabilities in any tasks. The standard practice that was adapted through the experiment's sessions was for the experimenters to reinforce this message at the start and end of each session. Given these strategies, this experiment was viewed as having low participant risks by both researchers and REB, enabling the experiments to commence.

At the conclusion of the final ballot in each meeting, a written post-meeting questionnaire was distributed for each participant to answer by themselves. At the conclusion of the third meeting'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.

Challenges to Ethical Design

During the experiment, we realized that participants were exposed to risks that were unaccounted for. As anticipated, some participants broke their promises to vote for a counterpart's preference. In some cases, the cheated participants became emotional and expressive. Expressions

where pointed, aggressive, and carried a lot of emotion, as exemplified by comments made by P31: "I thought we had a deal Jim", and P45: "you just lost all your credibility". In some instances, the cheated participant (ex. P12) held on to the emotions aroused by the cheater through multiple trials and tried to get back at the cheater. This behavior highlighted the fact that the emotions aroused from being cheated could be very strong. The strength of induced emotions can be attributed in part to the fact that wins and losses in the experiment were social phenomena, which can have an amplifying effect on the emotional payoff of winning or losing [6].

In three cases, cheating created significant tension in the meetings, compelling the facilitator to step in and try to defuse the tension in the virtual meeting room, contravening his instructions not to intrude in the group's discussions to avoid compromising the collected data. Although the facilitator did manage to defuse the tension in the meetings and calming down the participants, he was not trained in conflict resolution. In this case, the experimenter decided on the fly that defusing the tension and creating a calm experiment setting in which participants are more relaxed was preferable to a more tense scenario in which the participants were at risk of being subjected to psychological harm. Training in conflict resolution would have enabled the facilitator to better assess the meeting room and making a more measured decision on when the tension is too high and when it is low enough not to subject the participants to risk of social or psychological harm. The question of how much tension or social stress is appropriate in such experimental settings depends on the experiment and its potential benefits. We thus recommend all designers of group experiments to recognize that participants may get highly emotional in their social interactions and devise strategies to deal with resulting tension beforehand.

As mentioned earlier, the discussed ethical concern arose after the REB approval and during the experiments. Resolving these concerns would benefit more from more dynamic REB processes, as advocated by Munteanu et al. [12], in which the researchers not only have their experimental protocol reviewed, but also consult the board for recommendations about ethical concerns that rise when running the experiment.

References

- [1] 1978. The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research.
- [2] Nathan Bos, Karrie Karahalios, Marcela Musgrove-Chávez, Erika Shehan Poole, John Charles Thomas, and Sarita Yardi. 2009. Research ethics in the facebook era: privacy, anonymity, and oversight. In *CHI'09 Extended Abstracts*. ACM, 2767–2770.
- [3] Nathan Bos, Judy Olson, Darren Gergle, Gary Olson, and Zach Wright. 2002. Effects of four computer-mediated communications channels on trust development. In *Proceedings of the SIGCHI*. ACM, 135–140.
- [4] John K Butler. 1991. Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of management* 17, 3 (1991), 643–663.
- [5] Robyn M Dawes, Jeanne McTavish, and Harriet Shakerlee. 1977. Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of personality and social psychology* 35, 1 (1977), 1.
- [6] Daniel M. Doleys, Karen C. Wells, Steven A. Hobbs, Mark W. Roberts, and Lora M. Cartelli. 1976. The Effects of Social Punishment on Noncompliance: A Comparison with Timeout and Positive Practice. *Journal of Applied Behavior Analysis* 9, 4 (1976), 471–482.
- [7] Mark S Frankel and Sanyin Siang. 1999. Ethical and legal aspects of human subjects research on the Internet. *Published by AAAS online* (1999).
- [8] Jeffrey T Hancock, Jennifer Thom-Santelli, and Thompson Ritchie. 2004. Deception and design: The impact of communication technology on lying behavior. In *Proceedings of SIGCHI*. ACM, 129–134.
- [9] Craig Haney, W Curtis Banks, and Philip G Zimbardo. 1973. A study of prisoners and guards in a simulated prison. *Naval Research Reviews* 9, 1-17 (1973).
- [10] Allen M Hornblum. 2013. *Acres of skin: Human experiments at Holmesburg prison*. Routledge.
- [11] Lik Mui, Mojdeh Mohtashemi, and Ari Halberstadt. 2002. A computational model of trust and reputation. In *Proceedings of HICSS*. IEEE, 2431–2439.
- [12] Cosmin Munteanu, Heather Molyneaux, Wendy Moncur, Mario Romero, Susan O'Donnell, and John Vines. 2015. Situational Ethics: Re-thinking Approaches to Formal Ethics Requirements for Human-Computer Interaction. In *Proceedings of CHI*. ACM, 105–114.
- [13] Cosmin Munteanu, Heather Molyneaux, and Susan O'Donnell. 2014. Fieldwork with Vulnerable Populations. *Interactions* 21, 1 (Jan. 2014), 50–53.
- [14] Eric Noreen. 1988. The economics of ethics: A new perspective on agency theory. *Accounting, Organizations and Society* 13, 4 (1988), 359–369.
- [15] Willem Standaert, Steve Muylle, and Anirban Basu. 2013. Assessing the effectiveness of telepresence for business meetings. In *Proceedings of HICSS*. IEEE, 549–558.
- [16] Akbar Zaheer, Bill McEvily, and Vincenzo Perrone. 1998. Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization science* 9, 2 (1998), 141–159.