
Evaluating an Intuitive Teleoperation Platform Explored In a Long-distance Interview

Ritta Baddoura

Université Lyon 1 ; INSERM U846
18 avenue du Doyen Lépine,
69500 Bron, France
ritta.baddoura@inserm.fr

Gentiane Venture

Tokyo University of Agriculture and Technology
2-24-16 Nakacho Koganei-shi
184-8588 Tokyo, Japan
venture@cc.tuat.ac.jp

Guillaume Gibert

Université Lyon 1 ; INSERM U846
18 avenue du Doyen Lépine,
69500 Bron, France
guillaume.gibert@inserm.fr

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/authors.

***HAI '14 , Oct 29-31 2014, Tsukuba, Japan
ACM 978-1-4503-3035-0/14/10.
<http://dx.doi.org/10.1145/2658861.2658930>***

Abstract

SWoOZ is an intuitive teleoperation platform using a humanoid robot as a proxy between two humans: a remote user teleoperating the robot and a local user interacting directly with it. NAO (Aldebaran) is the proxy used in this study. The remote user controls its head motion with his own movements (live) while his real voice is transmitted to the local user with an unnoticeable lag time. This paper presents a user study of the platform in the context of a long-distance survey and investigates the possible effect of the remote user's previous experience with robots on the local users' evaluation of the proxy. Although found useful, likable and satisfying by all the local users, only the ones interviewed by the non-naïve user find it averagely credible. Results fail to validate an effect of the remote users' previous experience with robots on the local users.

Author Keywords

Teleoperation; user study; humanoid proxy; social robotics; previous experience with robots; head motion.

ACM Classification Keywords

Design; experimentation; human factors; measurement.

Introduction

Many interactions take place simultaneously when humans are communicating through a telerobotic system as reported in [6]. These interactions include Human-Robot Interaction (HRI) between the human local user and the remotely controlled robot, human-human interaction, and human-computer interaction between the remote user and the human partner's image on the screen. Various studies show the interest of using teleoperation [9, 11] and telepresence [5, 6, 7] robots in different fields such as remote education [12], health care environments or offices [7].

We developed SWoOZ, a teleoperation setup that mirrors face, eye and head motion on a humanoid robot and consequently allows the generation of spontaneous movements in order to support a natural interaction [3, 4]. In this setup, a humanoid robot is used as a proxy between two humans. One human, the remote user, is bound with the humanoid robot and controls its head motion in real-time and free of attached sensors. The remote user perceives the scene almost as if she or he was present in the same room with his human interlocutor, called the local user. The humanoid head motion for instance, as the local user sees it, is the direct translation of the remote user's motion which is accurately tracked and replicated by the robot with less than 200 ms delay. The SWoOZ platform can be used to manipulate specific movements of the robot without modifying the rest of the dynamics.

In order to evaluate this teleoperation platform when used in a realistic scenario, we started a study addressing different experimental questions. We present here the local users' evaluation of the performance of the SWoOZ teleoperated robot as well as examine the possible effect of the remote users'

previous experience with robotics (naive vs. non-naive) on these evaluations.

Methods

Experimental Setup & Equipment

The SWoOZ platform consists of:

- A system able to estimate the remote user's head pose (orientation and location) thanks to the Random Forests Head Tracking system [2] and to a consumer depth camera (ASUS/Xtion sensor).
- A software program to apply online manipulation to specific parameters.
- A humanoid robot: SWoOZ is compatible with NAO (Aldebaran) and iCub (<http://www.icub.org/>). NAO is used in the current study. NAO is a 58-cm tall humanoid robot. Its body has 25 degrees of freedom (DOF) whose key elements are electric motors and actuators. We used the programming software delivered with the robot to control it.

Once the data are estimated, they are sent to the robot so that it mimics the estimated motion of the remote user's head. Further information about the SWoOZ platform can be found in [3, 4] as well as on the SWoOZ Github page:

<https://github.com/GuillaumeGibert/swooz>.

The remote user's voice captured by a microphone is transmitted to the local user interacting with the teleoperated robot through a small speaker positioned behind it. An IMU (Inertial Measurement Unit) sensor is attached around the remote user's and the local user's heads for further motion analysis. To bind the remote user to the robot and enable her or him to sense the scene almost as if she or he was seated in its place, auditory and visual feedbacks are transmitted to her or

him using a High Definition (HD) webcam (Creative Live Cam Socialize HD) positioned behind the robot and binaural microphones (MS-TFB-2, The Sound Professionals, Inc.) discreetly placed on the robot's body.

Participants

Two male remote users (one naive and another non-naive) and 20 naive local users (previous exposure to robots was controlled prior to the experiment) volunteered to take part in the study. The 22 candidates are Japanese students from Tokyo University of Agriculture and Technology (TUAT). All of them range in age from 19 to 25 years old. The naive remote user (never used or interacted with robots before) interviewed 14 local users (9 males, 5 females). This group of local users will be referred as X in the rest of the text. The non-naive remote user (who has previously used NAO for HRI studies) interviewed 6 local users (5 males; 1 female). This group of local users will be referred as Y. Both were trained to perform the interview (e.g. learning the questions, keeping their behavior consistent with the experimental scenario, monitoring the answer duration of the local user) and were similarly instructed regarding the technical requirements necessary for the proper functioning of the setup. More generally, the experimental design aims at defining a precise and repeatable conversational context.

Materials, Procedure & Data collection

The experiment took place on a Japanese university campus. The local and the remote users volunteered to participate in an anonymous survey lead by a Japanese researcher working in France. They were informed about the following: a) The interview is live mediated by a humanoid robot, b) The interview room is filmed

using two cameras, c) The remote user's and the local user's voices are recorded and IMU sensors are used for head motion capture for future analysis.

The remote user teleoperating the robot is in room A while the robot NAO and the local user are facing each other in room B, a real meeting room (see **Figure 1**). When the local user is seated, the remote user addresses him or her, presents himself and provides a recapitulation of the survey context. The local user is reminded that there are no false and right answers: only personal opinions are expected. Then the interview starts. It consists of 15 questions revolving around the specificities of the Japanese and French cultures. The interview's overall duration is 10 min and when it is completed, the remote user asks the local user to fill the questionnaire placed on the table.

The questionnaire consists of 40 items divided into five sets. These sets assess various aspects of the local users' evaluation of the robot as a proxy that mediates their communication with the remote user. At this first phase of the study, only two sets of the questionnaire (the first and the fourth) are used to investigate the following: (1) the local users' evaluation of the robot proxy as teleoperated through the SWoOZ platform, (2) the effects of the remote user's previous exposure to robots (naive vs. non-naive user) on the local users' evaluation. These two sets use a 5-point Likert scale (the points go from 0 to 4 where 0 = not at all and 4 = to a very high degree). The questionnaire needs 5 to 10 min to be filled (depending on the local user). The interview and the questionnaire were carefully designed, tested and modified in accordance, following a pilot study on 4 local users in France and 3 local users in Japan.



Figure 1. The remote user is in room A while the local user and NAO are in room B.

The Cronbach's alpha of the questionnaire is 0.91 which is above the generally acceptable level 0.7 [10] and shows a very good internal reliability. Basic descriptive statistics as well as non-parametric tests (Mann-Whitney test and Spearman correlation rank) will be applied to the local users' answers to the questionnaire.

Hypotheses

We first hypothesize that the remote user's previous exposure to HRI will impact the local users' experience of the interaction as well as their ratings and that the evaluations made by X (interviewed by the naive remote user) will be significantly different from the ones made by Y (interviewed by the non-naive remote user) (H1).

Regarding the local users' evaluation of the humanoid proxy, the focus is on their ratings of its usefulness, likability, credibility and on their satisfaction with it. Based on a preliminary study performed using the SWoOz setup [4] as well as on the proxy's capacity to mirror the remote user's head motion and transmit his real human voice in real-time, we expect the local users' ratings of usefulness, likability and satisfaction to be above the average score, 2 being the average on this 5-point Likert scale (H2).

From another perspective, the authors in [8] showed that a robot's appearance affects its likability and that the participants expect the robot's appearance to match its task during an interview context. NAO is described as "*small, cute and round*" on Aldebaran's official website [1]. Knowing that the appearance of this 58-cm tall humanoid does not seem the most appropriate to match the image and the role of a researcher, we expect that the local users give poor ratings to its credibility (H3).

Results

We calculated the descriptive statistics (95% CI) based on the local users' ratings of the humanoid proxy's performance. X and Y generally gave it medium-to-low ratings and Y seemed to have given more generous scores than X.

X considered that the robot failed in being credible (X: M= 1.28, SD= 1.03) (H3 validated for X). Nevertheless, they found the proxy rather satisfying (X: M= 2.14, SD= 1.30), likable (X: M= 2.07, SD= 1.10) and useful (X: M= 2.5, SD= 1.11) (H2 validated for X). Y found the proxy averagely satisfying (Y: M= 2.00, SD= 0.81), likable (Y: M= 2.83, SD= 0.68) and useful (Y: M= 2.5, SD= 1.11) (H2 validated for Y). Contrary to X, they found the NAO proxy rather credible (Y: M= 2.33, SD= 1.10) (H3 infirmed for Y).

We did a Mann-Whitney test to ascertain if the differences between X and Y scores are statistically significant, thus implying an effect of the remote users' previous experience with robots on their evaluations of NAO's performance (see **Table 1**). The observed U-values ranged between 22 and 41.5 and failed to be significant as they were superior to the U critical value (17) at $p \leq 0.05$ (H1 infirmed).

Some interesting associations regarding the useful aspect of the proxy were revealed for both X and Y when calculating the Spearman rank correlations (see **Table 2**). These positive dependencies showed that the more the local users found NAO useful, the more they appreciated its credibility and the more they found it satisfying.

NAO is	Credible	Efficient	Engaging
U observed	22.0	41.5	32.0
	Useful	Likable	Satisfying
	32.0	31.0	38.0

Table 1. Mann-Whitney test results based on the evaluations made by X and Y of the NAO proxy (U critical = 17).

NAO is Useful and	X		Y	
Credible	corr.= 0.61	p = 0.01	corr.= 0.78	p = 0.06
Satisfying	corr.= 0.57	p = 0.03	corr.= 0.91	p = 0.01

Table 2. Spearman rank correlations based on the evaluations made by X and Y of the useful, the credible and the satisfying characters of the NAO proxy.

Discussion and Conclusion

The local users showed a moderate to low appraisal of the proxy's performance. The results failed to validate an effect of the remote users' previous experience with robots on their answers (H1 infirmed). The hypotheses 2 and 3 were validated for X but only hypothesis 2 was confirmed for Y who rated the proxy as averagely credible, against our expectations.

The validated correlations also showed that the proxy's credibility is positively correlated with its usefulness which is also associated to the satisfaction of the local users with it. Based on that and knowing that the proxy's usefulness was averagely rated, it is possible to suggest that in order to improve the local users' evaluation of the proxy's usefulness, one interesting possibility would be to work on enhancing its credibility by investigating what might have limited it in the present experience: the lack of adequacy of the robot's appearance with the researcher's role or its limited communication features. Indeed, it is possible that NAO's appearance and its interactive abilities did not match its given role and task as a mediator for distant inter-human communication. Would a robot with richer facial expressions or with a more android aspect, seem more credible to the local users, particularly to the ones interviewed by the naive remote user (X)?

When considering the difference between X and Y when rating the proxy's credibility, the question of whether the remote users' previous experience impacted their evaluations or not, appears legitimate. But the statistical results did not validate this effect. The small size of the sample interviewed by the non-naive user could explain this lack of significance. Thus, more exploration with a bigger sample is needed before completely ruling out a possible effect of the remote users' previous experience. Also, the intuitive mediation characteristics of the SWoOZ platform might have mitigated the remote users' effect. The proxy's mediation, as it interferes in the human dyad, is in favor of rendering a rather homogeneous behavior of the robot, especially that only the remote users' head motion is mirrored here. Using a humanoid robot that has richer facial expressions and that is able to

replicate more of the remote users' motion (e.g. eye, mouth or arms motion) might better transmit the differences existing between the remote users' previous experience with robots or with teleoperation.

The failure to prove a significant effect of the remote users' previous experience with robots is interesting in regards to the platform's usability as it strongly suggests that any naive person could, with some preparation, successfully use SWoOZ as a remote user and be as equivalently effective as a more experienced person in supporting the proxy's likable, useful and satisfying characters. The question of the proxy's choice, in relation to its appearance and to its assigned task, needs to be further investigated for credibility reasons. Knowing that the SWoOZ setup is compatible with different humanoid robots, the replication of our experiment with another humanoid would probably clarify our findings. This paper's encouraging results are to be completed with the analysis of the motion capture data. Obtained by testing the setup in a realistic communication scenario, they show that SWoOZ's intuitive and spontaneous teleoperation features draw stimulating perspectives for distant inter-human communication in the near future.

Acknowledgements

This work was supported by the ANR SWoOZ project (11PDOC01901) and the « Techno-Innovation Park » grant from Tokyo University of Agriculture and Technology. We thank Ryo Matsukata, Hiroshi Takaiwa, Manfred Corbeau, Takamune Izui from TUAT and Florian Lance from INSERM U846 for their help during the experiments.

References

- [1] Aldebaran Robotics.
<http://www.aldebaran.com/en/humanoid-robot/nao-robot>.
- [2] Fanelli, G., Gall, J. and Van Gool, L. Real time head pose estimation with random regression forests. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2011), 617-624.
- [3] Gibert, G., Petit, M., Lance, F., Pointeau, G. and Dominey, P.F. What makes human so different? Analysis of human-humanoid robot interaction with a super Wizard of Oz platform. In *Proc. IROS 2013, International Conference on Intelligent Robots and Systems* (2013).
- [4] Gibert, G., Lance, F., Petit, M., Pointeau, G. and Dominey, P.F. Damping robot's head movements affects human-robot interaction. In *Proc. of the 2014 ACM/IEEE HRI* (2014).
- [5] Harutyunyan, V., Manohar, V., Gezehei, I. and Crandall, J. W. Cognitive Telepresence in Human-Robot Interactions, *Journal of Human-Robot Interaction*, vol. 1 (2012), 158-182.
- [6] Kiselev, A. and Loutfi, A. Using a mental workload index as a measure of usability of a user interface for social robotic telepresence. *Workshop in Social Robotics Telepresence* (2012).
- [7] Kristoffersson, A., Coradeschi, S. and Loutfi, A. A review of mobile robotic telepresence, *Advances in Human-Computer Interaction*, 3 (2013).
- [8] Li, D., Rau, P.P. and Li, Y. A cross-cultural study: effect of robot appearance and task. *International Journal of Social Robotics* 2 (2010), 175-186.
- [9] Nishio, S., Ishiguro, H., Anderson, M. and Hagita, N. Representing Personal Presence with a Teleoperated Android: A Case Study with Family. In *AAAI Spring Symposium: Emotion, Personality and Social Behavior* (2008), 96-103.
- [10] Nunnally, J.C. *Psychometric theory*. McGraw-Hill (1978).
- [11] Riek, L.D. Wizard of Oz Studies in HRI: A systematic Review and New Reporting Guidelines. *Journal of Human-Robot Interaction* 1 (2012), 119-136.
- [12] Tanaka, F., Takahashi, T., Matsuzoe, S., Tazawa, N. and Morita, M. Telepresence robot helps children in communicating with teachers who speak a different language. In *Proc. of the 2014 ACM/IEEE HRI* (2014).