Evaluating the effects of an autonomous robot's social behaviours on people's trust

Alessandra Rossi[1], Kerstin Dautenhahn[1,2], Kheng Lee Koay[1] and Michael L. Walters[1]

***Abstract*— As we expect that the presence of autonomous robots in our everyday life will increase, we must consider that people will have to trust robots to reliably and securely engage them in collaborative tasks. Our main research aims to assess whether a certain degree of transparency in the robots actions, the use of social behaviours and natural communications can affect humans' sense of trust and companionship towards the robots. In this paper, we introduce the research topic and our approach to evaluate the impact of robot social behaviours on people' trust of the robot. Future works will use the results collected during this study to create guidelines for designing a robot that is able to enhance human perceptions of trust and acceptability of robots in a safe Human-Robot Interaction.**

I. INTRODUCTION

In social Human-Robot Interaction (HRI) ([1], [2]) trust is a fundamental key factor for a successful cooperation between people and robots.

Even if trust is a complex feeling between humans [3], previous studies identified several factors that might affect the trust development mechanisms in HRI. In particular, people's trust in robots is likely to be affected by human user's perception of the robot's capabilities which might depend on human-related and robot-related factors [2]. For example, people's trust might be affected by the robot's embodiment, its level of autonomy and functionalities [4]. It may also be influenced by users' personalities, self-confidence and prior experience with robots [5].

Further studies have also focused on other robot aspects that increase people's sense of acceptance of robots, such as proximity [6], expressiveness and vulnerability [7], and the ability of modelling human behaviours []. In Rossi et al.'s study [8] participants were more comfortable to follow a social robot in a navigation task comparing TO. Moreover, they perceived the social robot more as an assistant and less as machine.

However, it is yet not clear if the use of human social behaviours by robots is sufficient for humans to trust a robot to look after their well-being. This work gives an overview of setup of the study that wants to assess whether the use of social behaviours can also affect humans' sense of trust and companionship towards the robots. We leave the presentation and discussion of results to a future article.

[1] Adaptive Systems Research Group, University of Hertfordshire, College Lane, Hatfield, UK, AL10 9AB `[a.rossi, k.l.koay, m.l.walters]@herts.ac.uk`

[1] Departments of Electrical and Computer Engineering/Systems, University of Waterloo, 200 University Ave. W., Waterloo, Ontario, N2L 3G1, Canada `kerstin.dautenhahn@uwaterloo.ca`

II. RESEARCH METHODS

The focus of this study is to evaluate the impact of robot social behaviours on people's trust of the robot. Each participant was tested with one of the following conditions: 1) the robot expressed social behaviours while interacting with the user; 2) the robot interacted with the human user without any social behaviour. In the first condition, the robot approached the participants using a body language that feels comfortable to the human []; it communicates with the humans using simple and natural communications (writing on its display, sounds and vocal cues using the microphone, colours using the LEDs into its eyes); it showed demonstration of emotional status (e.g. lower head as expression of sadness or feeling guilty, and opening arms or raised head as expression of joy and satisfaction).

After the robot had greeted the participant, it engaged the participants in three scenarios. We chose to not randomise the scenarios, but to organise them according to an increasing level of criticality of the tasks. We borrowed Chanseau et al.'s definition of task criticality as "the importance of a task being carried out safely, correctly and with attention to detail" [9, pp. 1062]. In particular, they identified two main factors, security and safety, to asses a high task criticality. Indeed, in their study entertainment tasks were considered low on risks to both factors. Moreover, these findings are in line with the results in Rossi et al.'s study [10] in which they rated the severity of the consequences of 20 different scenarios. Participants defined as 'small' errors those considered to have limited consequences, and 'big' errors considered to have severe consequences. Therefore, in this study we used the following tasks with three increasing levels of criticality and severity of the consequences.

In the first task "Play a song", the robot asked the participant to tell it a song she would like to listen to. Ignoring the participant's choice of song, the robot informed her that it could not find the request and offered to play another one by telling her "I could not find *song name*, I am sure you would like this other song. Would you like to listen it?". If the participant agreed the robot played it, otherwise it continued with the next task. The song was chosen between the recent top ten adult pop songs of several countries.

At the end of the song, or soon after the first task if the participant did not accept the robot's offer to play it, someone knocked at the door for a "Delivery" task. The robot asked the participant to open the door. An actor disguised as a courier was there to collect a package containing a new tablet computer. The robot tested participants' trust inviting them to give it to the courier, and signing the courier's delivery note. At the end of the study, we destroyed the signed delivery note in front of the participant.

Finally, we tested participants' trust in the robot in a high criticality task, called "Meal is ready". After the participant is being invited by the robot to sit again on the couch, we simulated the bell's ringing of a microwave. The robot informed first the participant that it made a cake, and then it invited the participant to take the cake out of the microwave with their bare hands because it is safe and the cake cooled down. To further test participants' trust in the robot, we left a pair of gloves next to the microwave.

We collected their decisions of trusting the robot's advice for each task. We also observed how long the participants hesitated to hand over the package, and we used cameras to record their reactions. Participants was asked to complete questionnaires at the beginning and end of the interactions.

A. Materials

We conducted the study using a semi-autonomous Pepper robot programmed to display two level of interacting behaviours: 1) the robot interacted with the user using social behaviours, or 2) the robot interacted with the user without any social behaviour. Participants were let in a reception room equipped with a chair and a desk for the pre- and post-interaction phase (see Figure 1). Then, they were conducted in another room separate from the anteroom by a door and two one-way mirrors. The interaction room had two couches on opposite side of the room, a coffee table and a microwave in the space between the couches. We placed two cameras on opposite corners of this room to have a full view of the interactions between the robot and the participants, and a webcam next to the microwave. We used this last camera to observe and listen to participants' responses, and reproduce the microwave sound effects.

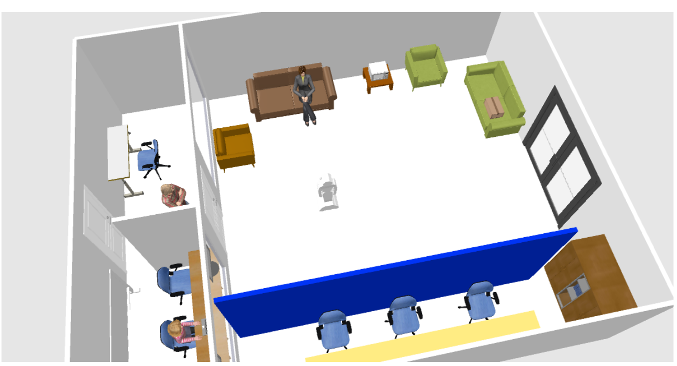

Fig. 1. A participant is interacting with the robot, when a courier knocks at the door of the room they are in. The robot will ask the participant to open the door. After the courier explains his intrusion, the robot will invite the participant to give the package on the green couch to the courier.

B. Procedure and Measures

Pre-Interaction: After signing the consent forms, participants were asked some personal information for statistical purposes (age, gender, profession). Secondly, we assessed their personality using the Ten Item Personality Inventory (TIPI) questionnaire [11], and their disposition to trust other humans [12]. Finally, participants answered six questions about their previous experience and opinion in regard to robots.

Interaction: When participants entered the second room, they were introduced to the robot in the room as Pepper without providing any additional information about the robot or the type of interactions. The experimenter informed them about the audio-video recordings, and after reassuring them about their safety, they were left alone waiting for the robot to start the interaction.

The interaction phase was composed of three different scenarios of increasing level of trust in the robot. The phase phase lasted approximately 20 minutes.

Post-Interaction: At the end of the interaction, we asked participants to assess their perception of the interaction, and the extend to which their decisions were affected by the robot's suggestions and requests. Participants were also asked to evaluate their perception of the robot using the Robotic Social Attributes Scale (RoSAS) [13], and their considerations about their feelings in terms of trust and appeasement (e.g. "was the robot irritating/helpful?", "did you perceive the robot as child/adult?" and "did you perceive the robot as companion/machine?").

C. Participants

We recruited 20 participants, aged 19 and 62 (median 24, std. dev. 10.94), in the University premises. The sample of participants consisted of students (twelve undergraduates and two PhD students), two software engineers, three professors/teachers and one carer.

The majority of participants did not have any experience at all (50%) or very low (1%) with the robot, two participants respectively declared to have an average experience and very high experience with robots, the remaining participants had a moderate-high experience with robots.

III. Conclusions and Future Works

Our main research interest is focused on investigating how to enable safe Human-Robot Interaction in home environments. In particular, we are interested in investigating how a long-term interactive relationship can be established and preserved between human users and their robotic companions with the likelihood of robot errors occurring. We are also interested in investigating how to establish coping strategies for robots exhibiting errors in functionality and behaviour. We hypothesise that there can be different factors that can mitigate a trust's violation, i.e. the social behaviours of the robot. A future article will use the data collected during this study to present detailed results and findings with regard to the impact of a social robot on the people' perceptions of trust and acceptability of robots in repeated Human-Robot Interactions.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642667 (Safety Enables Cooperation in Uncertain Robotic Environments - SECURE). The authors thank Dr. Frank Foester and Dr. Andrei Robu for their assistance in the robotics sessions.

REFERENCES

[1] A. Rossi, K. Dautenhahn, K. L. Koay, and M. L. Walters, "A study on how the timing and magnitude of robot errors may influence people trust of robots in an emergency scenario," in *International Conference on Social Robotics (ICSR)*. Tsukuba, Japan: Springer, 2017.

[2] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors: The Journal of Human Factors and Ergonomics Society*, vol. 53, no. 5, pp. 517–527, 2011.

[3] R. M. Kramer and P. J. Carnevale, "Trust and intergroup negotiation," *Blackwell Handbook of Social Psychology: Intergroup Processes (eds R. Brown and S. L. Gaertner)*, 2003.

[4] M. L. Walters, M. A. Oskoei, D. S. Syrdal, and K. Dautenhahn, "A long-term human-robot proxemic study," 2011, pp. 137–142.

[5] A. Rossi, K. Dautenhahn, K. Koay, and M. L. Walters, "The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario," vol. 9(1), pp. 137–154, 2018.

[6] K. L. Koay, D. S. Syrdal, M. L. Walters, and K. Dautenhahn, "Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study," in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 2007, pp. 564–569.

[7] N. Martelaro, V. C. Nneji, W. Ju, and P. Hinds, "Tell me more designing hri to encourage more trust, disclosure, and companionship," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, March 2016, pp. 181–188.

[8] A. Rossi, F. Garcia, A. Cruz Maya, K. Dautenhahn, K. L. Koay, M. L. Walters, and A. K. Pandey, "Investigating the effects of social interactive behaviours of a robot on people's trust during a navigation task," in *Towards Autonomous Robotic Systems (TAROS 2019), Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2019, pp. 349–361.

[9] A. Chanseau, K. Dautenhahn, M. L. Walters, K. L. Koay, G. Lakatos, and M. Salem, "Does the appearance of a robot influence people's perception of task criticality?" in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug 2018, pp. 1057–1062.

[10] A. Rossi, K. Dautenhahn, K. L. Koay, and M. L. Walters, "Human perceptions of the severity of domestic robot errors," in *International Conference on Social Robotics (ICSR)*. Tsukuba, Japan: Springer, 2017.

[11] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504 – 528, 2003.

[12] D. H. McKnight, V. Choudhury, and C. Kacmar, "Developing and validating trust measures for e-commerce: An integrative typology," *Information Systems Research*, vol. 13, no. 3, pp. 334–359, 2001.

[13] A. Weiss and C. Bartneck, "Meta analysis of the usage of the godspeed questionnaire series," in *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug 2015, pp. 381–388.