Getting to know Pepper

Effects of people's awareness of a robot's capabilities on their trust in the robot

Alessandra Rossi
Adaptive Systems Research Group,
University of Hertfordshire
Hatfield, UK
a.rossi@herts.ac.uk

Patrick Holthaus
Adaptive Systems Research Group,
University of Hertfordshire
Hatfield, UK
p.holthaus@herts.ac.uk

Kerstin Dautenhahn
Departments of Electrical and
Computer Engineering/Systems,
University of Waterloo
Waterloo, Ontario, Canada
Adaptive Systems Research Group,
University of Hertfordshire
Hatfield, UK
kerstin.dautenhahn@uwaterloo.ca

Kheng Lee Koay
Adaptive Systems Research Group,
University of Hertfordshire
Hatfield, UK
k.l.koay@herts.ac.uk

Michael L. Walters
Adaptive Systems Research Group,
University of Hertfordshire
Hatfield, UK
m.l.walters@herts.ac.uk

ABSTRACT

This work investigates how human awareness about a social robot's capabilities is related to trusting this robot to handle different tasks. We present a user study that relates knowledge on different quality levels to participant's ratings of trust. Secondary school pupils were asked to rate their trust in the robot after three types of exposures: a video demonstration, a live interaction, and a programming task. The study revealed that the pupils' trust is positively affected across different domains after each session, indicating that human users trust a robot more the more awareness about the robot they have.

CCS CONCEPTS

• **Computer systems organization → Robotics**; • **Computing methodologies → Cognitive robotics**; • **Human-centered computing → User studies**;

KEYWORDS

Human-Robot Interaction, Trust in robots, HRI awareness, Social robotics, UK robotics week

ACM Reference Format:

Alessandra Rossi, Patrick Holthaus, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters. 2018. Getting to know Pepper: Effects of people's awareness of a robot's capabilities on their trust in the robot. In *6th International Conference on Human-Agent Interaction (HAI '18), December 15–18, 2018, Southampton, United Kingdom.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3284432.3284464

HAI '18, December 15–18, 2018, Southampton, United Kingdom

ACM ISBN 978-1-4503-5953-5/18/12…$15.00
https://doi.org/10.1145/3284432.3284464

1 INTRODUCTION

Trust is widely assumed to be one of the key factors in human users' acceptance of social robots in human-centred environments [24]. However, a human user's awareness of the robot's skills also has significant effects on the interaction quality [1]. This work hence investigates how human trust in a social robot is affected by their interaction history and the human's knowledge about the robot's capabilities and limitations.

Trust between humans is constructed from a perception of ability, benevolence and integrity [25]. In Human Computer Interaction, Muir and Moray [26] showed that people's trust in a machine was strongly affected by the machine's good performance. Indeed, trust is a key factor in the acceptance of an autonomous robot as a peer, assistant or companion in human-centred environments. It can determine humans' perception of the usefulness of imparted information and capabilities of a robot [18, 31, 32].

Human awareness of a social robot's skills can be gained through robot appearance and behaviours including their common interaction history [19]. Typically, people naive to social robots in terms of real world encounters, already have certain expectations on their functionalities based on fictional movies and stories. In reality though, there is a significant gap between the current state of robotics research and science fiction [20], and sometimes even advertisements for real robots that make use of artificial intelligence[1]. As a consequence, negative effects on the interaction quality have to be considered when violating the user's expectations on the robot [23].

Within this paper, we investigate the relationship between human users' expectations of the robot and the quality of a Human-Robot Interaction (HRI). Particularly, we analyse the impact of repeated interactions that reveal different aspects of the robot's capabilities step-by-step on the users' trust ratings of the robot. With this approach, we gain insights on how human awareness of the robot affects their trust in it.

[1] https://www.youtube.com/watch?v=SSecbMFQK1I

We hypothesise that human attitudes towards robots change when they become aware of the real potential and limitations of the robot. With our findings, we improve the way an interactive relationship between human users and their robotic companion can be established on different knowledge levels about the robot.

A user study has been carried out to gain insights as part of an event that introduces the general public to robotics research. Participants were introduced to different aspects of the robot in three sessions: They first saw a commercial about the robot, then were able to interact freely with it, and finally they programmed and tested social behaviours for a storyteller robot. After each session, we asked them to rate their trust in the robot using similar questionnaires. The study supports our hypothesis and reveals that trust towards the robot is affected positively the more insights participants gain about the robot. Furthermore, trust alters similarly across different domains and qualities such as cognitive tasks or assistance in cases of danger.

The next section introduces the phenomena of trust and human awareness in HRI. Afterwards, Section 3 describes the conducted study in greater detail. In Section 4, we present the evaluation of questionnaires. Results are presented and discussed in groups that asses the robot's companionship qualities, its trustability, as well as programming-related questions. We conclude and summarize our contribution in Section 5 and give a short outlook on future work.

2 BACKGROUND

In this Section, we provide an overview of the current state-of-art in HRI, introducing the development of trust in human-robot interactions, the effects on HRI awareness on people's perception of robots, and the use of emotions to improve an interaction between human users and robots.

2.1 Trust in Robots

As trust is a widely researched phenomenon in Human-Human, Human-Computer and Human-Robot Interactions, there are different definitions of trust. In this study we adopted one of the most prevalent definitions [39]: "Trust can be defined as the attitude that an agent will help achieve an individual's goals in a situation characterised by uncertainty and vulnerability" [21].

In Human-Computer Interaction, Muir and Moray [26] showed that peoples' trust in a machine was strongly affected by the machine's good performance. Indeed, trust is a key factor in the acceptance of an autonomous robot as a peer, assistant, or companion in human-centred environments. It can determine humans' perceptions of the usefulness of imparted information and capabilities of a robot [18, 31, 32]. Trust is constructed from a perception of ability, benevolence and integrity [25]. Higher trust is associated with the perception of higher reliability [30]. Moreover, other aspects can affect the peoples' perceptions of robots, such as the embodiment of a robot [22, 24, 33], and a deeper awareness of a robot's functionalities and operation.

Bainbridge et al. [2] found that participants were happy to follow a robot's instructions to throw books in the trash if the robot was present in the room with them, but not when the robot was not physically in the same room. In Walters et al.'s [38] work, people generally preferred robots with more human-like appearance and attributes but their personal characteristics changed their perception of the robot. For example, introvert people and people with lower emotional stability tended to prefer a robot with a more mechanical appearance. In a preliminary study, Wainer [37] compared the effects of different type of embodiments (a physical robot, a simulated one, a presence through a co-located robot and a remote tele-present robot) on people' perceptions of social interactions. They noticed that people favoured physically embodied interactions over both virtual and remote tele-conference interactions.

Yu et al. [39] investigated the correlation between a user's reliance on a system and their trust level. They showed that participants formed their judgements at the beginning of interaction and eventually adjusted it later on, depending on the systems performance. Rossi et al. [32] showed that the perceived trust of the human in the robot drops drastically when a robot presents behaviours that can lead to severe consequences.

2.2 Human awareness in Human-Robot Interaction

In Human-Human interactions, situational awareness can change the way peoples interact with each other, as well as it might change the outcomes of the interaction. For example, people aware of a disease in their proximity take precautions that can reduce the spread of the disease for themselves and others [16], e.g. avoiding contacts with the infected person, wearing protective masks, vaccinating. Marketing advertisements involve the use of persuasion techniques to induce people to buy a product or accept better a communication. People's persuasion knowledge can shape how they respond to the persuasion attempt [14].

In Human-Robot Interaction, Atkinson et al. [1] suggested that humans' trust in robots, and consequentially a successful interaction, increases proportionally to a greater shared awareness of the agents involved, activities, and situations between human users and robots. Tseng et al. [36] developed a human awareness Decision Network model where the robot adapts its behaviours in respond to the different feedback of the user for meeting her expectation.

Moreover, a lack of human awareness in robots might lead the person to overtrust the robot and its functionalities. Abney et al. [27] define overtrust as the willingness of a person to accept the risk of delegating a task to a robot if 1) she believes that the robot is able to complete it or 2) her expectation is that the robot is able to mitigate the risk. Borenstein et al. [4] found that 62% of pediatric patients, their parents, and other caregivers would trust a robotic exoskeletons to be able to handle dangerous situations even if the robot did not have that capability. Booth et al. [3] investigated participants' responses to a robot's request to move in a secure-access student dormitory. They conducted the experiment with two conditions: 1) an anonymous robot and 2) a food delivery robot, where both asked to enter the building. They observed that participants were more likely to let the food delivery robot enter the building or in situations when they were in a group.

In Human-Robot Interaction several definitions of "human awareness" exist [9, 10, 12, 17]. In this work, we define HRI awareness simply as "the human understanding of the capabilities and functionality of a robot: to be aware of environment and people presence around itself; to interact with one or more humans according social conventions; to perform a specific activity; and to have artificial intelligence."

2.3 Emotions in Human-Robot Interaction

Emotions are natural human behaviours and "they are parts of the very process of interacting with the environment."[15, page 51]. Frijda et al. also propose that emotions have direct or indirect social consequences on the individuals involved in the HRI. Therefore, exhibition of socially interactive behaviours is a key factor in human acceptance of an autonomous robot [6, 7].

In particular, a robot that is able to show emotions helps to facilitate the interactions with a human [5]. Ficocelli et al. [13] defined a model to determine the appropriate emotions that a robot needs to show to elicit the well-being of a patient in an assistive interaction. Rincon et al. [29] presents a robot that was able to have non-verbal communications, like perceiving and displaying human emotions, to elicit empathy in a daycare center.

Syrdal et al. [35] showed that a Pioneer robot with dog-inspired affective cues communicates a sense of affinity and relationship with humans. Song et al. [34] used a robot that was able to express emotions through colours, sounds and vibration to solicit a more natural interaction between people and robots.

Humans can feel several emotions however, Ekman [11] described in his work six universally recognised basic emotions: anger, disgust, fear, happiness, sadness and surprise. We chose the emotions, we believed were connected to the story presented to the participants, inspired by Ekman's [11] findings.

Emotions are not the main focus of this study, but we used them to provide a pleasant and interesting experience in a robot storyteller scenario.

3 APPROACH

The UK Robotics Week[2] provides an annual opportunity to help students in the United Kingdoms's schools, colleges and universities to take an interest in modern technologies. Within this context, we decided to present a series of events to introduce school pupils to the state-of-art of social Human-Robot Interaction (HRI) and some currently adopted social cues.

The pupils were exposed to three different types of HRI: a video HRI, a real live HRI and HRI programming of a robot. In particular, during the programming sessions, pupils were focused on the implementation of emotions in HRI. We collected pupils' perceptions of the robot through questionnaires.

The focus of the event was both on the design of emotions for the interactive humanoid robot called Pepper using body movements, gesture and other non-verbal cues, and on collecting pupils' perceptions of the robot during three different levels of interactions by means of questionnaires.

3.1 Method

We observed and analysed participants' behaviours during three different levels of interactions with the SoftBank Robotics robot Pepper. Pupils worked with two Pepper robots for one session that lasted approximately three hours including a short break. Participants were asked to watch a video of the Pepper robot, then they were presented with two real Pepper robots, which they programmed and then tested different behaviours for the robots. In order to analyse the interactions between the human participants and the robot, we asked the participants to answer a questionnaire at the end of each of activity. All participants received a certificate of participation at the event.

3.2 Procedure

The event is organised in three different stages: 1) meeting the humanoid robot Pepper, 2) interacting with Pepper, and 3) programming Pepper. In the first part of the event, pupils watched a brief introductory video in which an actor interacted with the robot.

The second part of the event was focused in a closer interaction with Pepper in which pupils can touch the robot and use its built-in awareness function[3] and the "tickle me" scenario[4] (see Figure 1).

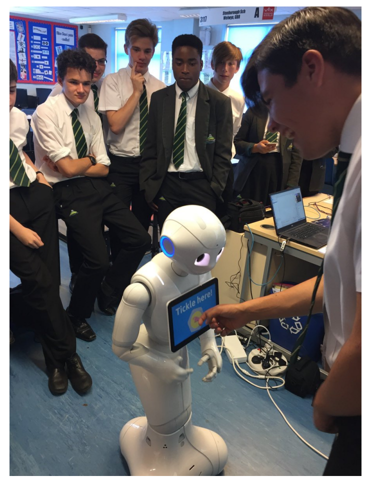

Figure 1: Two pupils are interacting with the robot during the second stage of the study.

Finally, participants built a story combining a given narrative with self-defined behaviours. The story consisted of a simplified version of the Hansel and Gretel fairy tale, composed of six different sentences conveying one of the six basic emotions (cf. Section 2.3):

- fear: Hensel and Gretel were rushing into the deep dark wood.
- anger: Gretel clenched her teeth firmly because of the way their parents treated them.
- sadness: But Hensel was feeling lonely and missed their family.
- surprise: Suddenly, an old lady appeared out of nowhere!
- disgust: She had an ugly green wart on her nose and smelled strangely.
- happiness: Fortunately, they followed their intuition, decided to return home and reunited with their parents.

The custom behaviours could be designed and implemented in the robot's graphical programming suite Choregraphe[5]. To achieve this goal, the pupils could use predefined building blocks that allowed

[2] http://hamlyn.doc.ic.ac.uk/uk-ras/robotics-week-2018

[3] http://doc.aldebaran.com/2-4/family/pepper_user_guide/life_pep.html#alife-pep

[4] http://doc.aldebaran.com/2-4/getting_started/samples/sample_interactive.html

[5] https://community.ald.softbankrobotics.com/en/resources/faq/developer/what-choregraphe

them to manipulate body movements, gesture, verbal and non-verbal cues individually per sentence and thus emotion.

In order to analyse the pupils' interactions and their experiences with the robot, the pupils were presented with a short questionnaire at the end of each stage.

We asked pupils different questions at the end of each part of the interaction about: 1) previous experiences with Pepper; 2) assessing participants' willingness to have Pepper as companion; 3) their perception of trust of Pepper; 4) their experience in programming the robot.

The questions sets were:

- Q1 Have you ever seen Pepper before today?
- Q2 If yes, where:
- Q3 Would you like to have Pepper in your home?
- Q4 Do you trust Pepper to be able to help you with your homework?
- Q5 Do you trust Pepper to wake you up in time for going to school?
- Q6 Do you trust Pepper to be able to warn you of danger, e.g. when using the Internet?
- Q7 Do you trust Pepper to help you in case of danger?
- Q8 Programming Pepper was? [very boring - very fun]
- Q9 Programming Pepper was? [very hard - very easy]
- Q10 Would you like to program Pepper again?

Questions Q1 and Q2 were asked only after the video interaction, while pupils answered the questions Q8-Q10 only at the end of the programming phase. We repeated the sets of questions Q3-Q7 after all the interactions. We chose these simple scenarios the pupils are familiar with. Indeed, since primary school, British children are taught about the danger of internet, such as cyberbullying or fishing. They know that is important to wake up in time to go to school and doing properly their homework.

3.3 Participants

We conducted the event in a local secondary school. Participants were secondary school pupils in year groups 10 to 12 [min age 14, max 15, avg. age 14.76, std. dev. 0.42]. The event was conducted over two consecutive days in the school and participants were tested in their age year groups, making a grand total of 43 pupils [6 girls, 37 boys].

4 RESULTS

As part of the questionnaire, we were interested in participants' previous experiences, their perceptions and expectations towards the Pepper robot. The majority of participants (86%) declared to not have any previous experience with this robot, while one participant was not sure if she saw it in a program TV. Participants with previous experiences with Pepper were: 1) watching a TV program (e.g a documentary at BBC); 2) live show about humanoid robotics; advertisement 3) on a poster, 4) on Snapchat, 5) on the web.

4.1 Questions Q3: Companionship

Participants expressed their willingness of having Pepper in their home through a Yes/No/Maybe measure. The results of our study show that participants were uncertain about Pepper's likeability after the video only. The majority of participants (58.13%) is not sure if they would like to have Pepper in their home. 27.9% would have liked to have Pepper in their home and only 13.95% would have not preferred to have the robot. Contrarily, the majority of participants (76.75%) expressed the preference of having Pepper in their home after the live HRI, a preference that decreased (60.46%) after the programming HRI. After the live HRI, 18.6% pupils showed uncertainty about the question, and only two pupils still did not want Pepper in their home. After the programming HRI, 25.58% participants were uncertain in their willingness of having Pepper in their home, and the remaining expressed negative consent.

Table 1: Paired samples t-test analysis comparing the means of participants' responses at Question Q3 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q3 (pepper in home)	t	p	95% CID
video - live	$t\_42 = -6.496$	$p < 0.001$	$-0.762 - -0.401$
video - programming	$t\_42 = -8.827$	$p < 0.001$	$-1.057 - -0.664$
live - programming	$t\_42 = -3.334$	$p < 0.001$	$-0.448 - -0.110$

We also found a statistically significant correlation between the willingness of having the Pepper robot in their home and the effects of the interaction ($p(2, 84) < 0.001, F = 47.162$). We performed a t-test analysis on the interactions' paired samples and we found that there are significant average differences between the participants' perceptions of the robot and the type of interaction. On average, the ratings were higher after the live HRI than the video HRI, and they were higher after the programming HRI than both video and live HRI (see Table 1).

Discussion

We hypothesise that the participants' perceptions of the robot changed so drastically from the first interaction due for two reasons. We showed participants a commercial video of Pepper and the majority of participants did not have any previous experience with the robot. People are acquainted to advertisements, and when they do not know the brand, they can have a higher critical judgement on the advertising value [28]. On the other hand, the awareness of real capabilities of the robot, acquired during the programming phase, mitigated the negative participants' perceptions of the interactive video interaction with the real robot.

4.2 Questions Q4-Q7: Trust in the robot

Participants answered questions Q4-Q7 using a 5-point Semantic Differential Scales where 1 corresponds to "definitely no" and 5 corresponds to "definitely yes". All the ratings with values less then 3 were categorised as negative response, with values equal to 3 were considered uncertainty and with values more than 3 were categorised as positive responses.

Question Q4 (homework). Similarly at question Q3, we observed an increase of participants' trust in the robot after the live interaction and a decrease of their trust after the programming interaction.

Table 2: Paired samples t-test analysis comparing the means of participants' responses at Question Q4 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q4 (homework)	t	p	95% CID
video - live	$t\_42 = -3.334$	$p < 0.001$	$-0.762 - -0.401$
video - programming	$t\_42 = 8.323$	$p < 0.001$	$0.969 - 1.589$
live - programming	$t\_42 = 12.943$	$p < 0.001$	$1.4333 - 1.962$

After the video HRI, the majority of participants (74.42%) would trust Pepper to be able to help with their homework, and the 18.6% were uncertain in the robot's ability of perform such task. After the live HRI, the majority of participants that trusted Pepper increased at 88.37% while uncertain participants decreased at 9.3%. After the last interaction, we observed a decrease of participants that trusted the robot's capability of performing the task (58.14%) and an increase of participants who were not completely confident in the robot (32.56%).

We observed statistically significant correlation between participants' trust in the robot to be able to do participants' homework and the different types of interaction ($p(2, 84) < 0.001, F = 82, 949$). On average, the significant differences can be observed between the interactions (see Table 2). In particular, participants trusted the robot more after the live interaction comparing video and live HRIs, and live and programming HRIs. However, the participants' trust in the robot is higher after the programming HRI, comparing video and programming HRIs.

Question Q5 (alarm clock). We asked participants if they trusted robot to be able to wake them up for going to school. We observed that participants were divided between trusting the robot (55.81%) and not confident in their trust in the robot (34.88%) after the video interaction. Again, the live interaction with a real robot increased their trust in the robot (79.07%) and only the 16.27% were unsure if the robot was able to complete the task. The pupils' perceptions of the robot's capabilities changed after the programming session, 67.44% of participants declared to trust the robot with the task, while 25.58% remained uncertain.

We found a statistically significant correlation between participants' trust in the robot to be able to wake them up for going to school and the different HRIs they were tested with. ($p(2, 84) = 0.02, F = 4.097$). In particular, participants rated higher their trust in the robot after the live HRI than the video HRI (see Table 3).

Question Q6 (danger warning). We observed that participants were sceptical about the robot's ability of warn them of a danger after the interactions. Indeed, the majority of participants (46.51%, 44.18% and 51.16% respectively after the video HRI, live HRI and programming HRI) were not sure to trust the robot. Their trust in the robot (37.21%) increased slightly only after the live interaction (41.86%) and decreased again after the programming interaction (27.90%).

Table 3: Paired samples t-test analysis comparing the means of participants' responses at Question Q5 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q5 (alarm clock)	t	p	95% CID
video - live	$t\_42 = -3.223$	$p = 0.002$	$-0.681 - -0.156$

Table 4: Paired samples t-test analysis comparing the means of participants' responses at Question Q6 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q6 (danger warning)	t	p	95% CID
video - programming	$t\_42 = -2.622$	$p = 0.01$	$-0.905 - -0.118$

Table 5: Paired samples t-test analysis comparing the means of participants' responses at Question Q7 according the three different HRIs. For each pair of interactions it shows the t values, p-value corresponding to the given test statistic t, the upper and lower bounds of the confidence interval.

Q7 (danger help)	t	p	95% CID
video - programming	$t\_42 = -2.973$	$p = 0.005$	$-0.820 - -0.157$
live - programming	$t\_42 = -2.673$	$p = 0.01$	$-0.816 - -0.114$

Participants' trust in the robot to be able to warn them of a danger was found positively correlated with the three different HRIs ($p(2, 84) = 0.009, F = 5.014$). We also found that participants trusted the robot more after the programming HRI (see Table 4).

Question Q7 (danger help). Similarly to question Q6, participants were not completely confident in the robot's capability of helping them in case of danger after the interactions. After the video HRI, participants' were divided between not trusting the robot (41.86%) and not being confident in trusting the robot (44.19%). Interestingly, while after the live HRI this division was equal (41.9%), after the programming HRI 11.67% participants declared to trust the robot to be able to help them in case of danger and 53.5% participants remained uncertain.

We also observed a statistically significant correlation between the pupils' trust in the robot's capability of helping them in case of danger and the HRIs they were tested with ($p(2, 84) = 0.003, F = 6.211$). On average, participants trusted the robot with the task more after the programming HRI comparing both the video and live interactions (see Table 5).

Discussion

The results of this study show that participants' perceptions of trust in the robot were partially affected by the awareness of the robots' capabilities. We hypothesise that their trust in the robot might have been affected also by the embodiment of the robot. Indeed, when we asked the participants to rate their trust in the robot in waking them to go to school, the live interaction had a greater effect on their perception of trust than the video interaction.

Moreover, we observed that a higher participants' awareness of the robots' capabilities affected the participants' trust in the robot to be able to detect and handle dangerous scenarios. This is not a surprisingly result. According to Deutsch [8], risk-taking and trusting behaviour are different sides of the same coin, and a person is willing to take a risk only if the odds of a possible positive outcome are greater than those for a potential loss. Indeed, participants trusted the robot more if they know the real potential and limitations of the robot. They believe to be able to program the robot to handle the dangerous situations themselves, or have someone else available to program it for them.

4.3 Questions Q8-Q10: Programming Pepper

At the end of the programming interaction, we asked participants to rate their experience of programming the Pepper robot.

Participants answered questions Q8-Q10 using a 5-point Semantic Differential Scales where 1 corresponds respectively to "very boring", "very hard" or "definitely no" and 5 corresponds to respectively "very fun", "very easy" or "definitely yes". All the ratings with values less then 3 were categorised as negative response, with values equal to 3 were considered moderate and with values more than 3 were categorised as positive responses.

Question Q8 (fun). The majority of pupils (74.42%) thought that programming the robot was fun, while only two participants did not enjoy the experience. In the comment section they declared that they did not understand how to program the robot.

Question Q9 (simplicity). At the question how easy it was for them to program the robot, 62.8% expressed a moderate response and 20.93% declared it was easy to program the robot. The remaining participants believed it was very hard to program the robot.

Question Q10 (recurrence). Pupils gave a moderate or positive feedback. Indeed 23% of participants would like to program the robot again, while only two percent prefer to not repeat the experience. We observed from the open-ended questions that one of the two participants considered the robot terrifying, the other did not understand the programming explanations.

Discussion

As a public engagement activity, the event was a success for the pupils and they enjoyed programming the robot. However, partially disguised by the teachers, we underestimated participants' previous knowledge in programming and the time required to program emotional behaviours in the limited time we have been granted by the school. For future investigations, we will consider a more exhaustive pre-session to further explain participants how to use the program Choreographe and program different behaviours for the robot.

5 CONCLUSION

With the event, we successfully familiarized pupils with scientific field of social robotics. The majority of participants stated they were happy about their interactions with the robots and the programming activities. However, some pupils stated that programming the robot was not an easy task.

Our study supports our hypothesis and shows that participants' awareness of the robots' real potential and limitations affected their perceptions of trust in the robot. Participants' awareness of being able to program different robots' behaviours led them to believe that the robot is able to handle critical situations and cognitive tasks, such as helping them with their homework.

Furthermore we conclude that their trust in the robot might have been affected also by the embodiment of the robot. Indeed, we observed that the live interaction with the robot affected the participants' willingness of having Pepper in their homes increasing the effect.

Future works will integrate these findings to investigate how an interactive relationship can be established and preserved between human users and their robotic companions in short-term and long-term interactions.

Acknowledgments. 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).

REFERENCES

[1] D. J. Atkinson, W. J. Clancey, , and M. H. Clark. 2014. Shared awareness, autonomy and trust in human-robot teamwork. In *In Artificial Intelligence and Human-Computer Interaction: Papers from the 2014 AAAI Spring Symposium on.*

[2] W. A. Bainbridge, J. W. Hart, E. S. Kim, and B. Scassellati. 2011. The Benefits of Interactions with Physically Present Robots over Video–Displayed Agents. *International Journal of Social Robotics* 3, 1 (2011), 41–52.

[3] Serena Booth, James Tompkin, Hanspeter Pfister, Jim Waldo, Krzysztof Gajos, and Radhika Nagpal. 2017. Piggybacking Robots: Human-Robot Overtrust in University Dormitory Security. ACM, 426–434.

[4] J. Borenstein, A. R. Wagner, and A. Howard. 2018. Overtrust of Pediatric Health-Care Robots: A Preliminary Survey of Parent Perspectives. *IEEE Robotics Automation Magazine* 25, 1 (March 2018), 46–54. https://doi.org/10.1109/MRA.2017.2778743

[5] Jessica Rebecca Cauchard, Kevin Y. Zhai, Marco Spadafora, and James A. Landay. 2016. Emotion Encoding in Human-Drone Interaction. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI '16).* IEEE Press, Piscataway, NJ, USA, 263–270.

[6] Kerstin Dautenhahn. 2007. Methodology & themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems* 4, 1 SPEC. ISS. (2007), 103–108. https://doi.org/10.5772/5702

[7] Kerstin Dautenhahn. 2007. Socially intelligent robots: dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, 1480 (2007), 679–704.

[8] M. Deutsch. 1958. Trust and suspicion. In *The Journal of Conflict Resolution*, Vol. 2(4). 265âĂŞ279. https://doi.org/10.1177/002200275800200401

[9] Jill Drury. 2000. Extending Usability Inspection Techniques for Collaborative Systems. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00).* ACM, New York, NY, USA, 81–82. https://doi.org/10.1145/633292.633341

[10] J. L. Drury, J. Scholtz, and H. A. Yanco. 2003. Awareness in human-robot interactions. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, Vol. 1. 912–918. https://doi.org/10.1109/ICSMC.2003.1243931

[11] Paul Ekman. 1999. *Basic Emotions.* T. Dalgleish and M. Power (Eds.). 45–60 pages.

[12] Mica R. Endsley. 1988. Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting* 32, 2 (1988), 97–101. https://doi.org/10.1177/154193128803200221

[13] M. Ficocelli, J. Terao, and G. Nejat. 2016. Promoting Interactions Between Humans and Robots Using Robotic Emotional Behavior. *IEEE Transactions on Cybernetics* 46, 12 (Dec 2016), 2911–2923. https://doi.org/10.1109/TCYB.2015.2492999

[14] Marian Friestad and Peter Wright. 1994. The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research* 21, 1 (1994), 1–31.

[15] N. H. Frijda and B. Mesquita. 1994. *The social roles and functions of emotions.* Washington, DC: American Psychological Association. 51–87 pages.

[16] Sebastian Funk, Erez Gilad, Chris Watkins, and Vincent A. A. Jansen. 2009. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences* 106, 16 (2009), 6872–6877. https://doi.org/10.1073/pnas.0810762106

[17] Carl Gutwin, Gwen Stark, and Saul Greenberg. 1995. Support for Workspace Awareness in Educational Groupware. In *The First International Conference on Computer Support for Collaborative Learning (CSCL '95).* L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 147–156. https://doi.org/10.3115/222020.222126

[18] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y. C. Chen, Ewart J. de Visser, and Raja Parasuraman. 2011. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors: The Journal of Human Factors and Ergonomics Society* 53, 5 (2011), 517–527.

[19] F. Hegel, S. Gieselmann, A. Peters, P. Holthaus, and B. Wrede. 2011. Towards a typology of meaningful signals and cues in social robotics. In *2011 RO-MAN.* 72–78. https://doi.org/10.1109/ROMAN.2011.6005246

[20] Sarah Kriz, Toni D Ferro, Pallavi Damera, and John R Porter. 2010. Fictional robots as a data source in HRI research: Exploring the link between science fiction and interactional expectations. In *RO-MAN, 2010 IEEE.* IEEE, 458–463.

[21] J. D. Lee and K. A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, 1 (2004), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

[22] M. Ligthart and K. P. Truong. 2015. Selecting the right robot: Influence of user attitude, robot sociability and embodiment on user preferences. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* 682–687. https://doi.org/10.1109/ROMAN.2015.7333598

[23] Manja Lohse. 2009. The role of expectations in HRI. *New Frontiers in Human-Robot Interaction* (2009).

[24] N. Martelaro, V. C. Nneji, W. Ju, and P. Hinds. 2016. Tell me more designing HRI to encourage more trust, disclosure, and companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* 181–188. https://doi.org/10.1109/HRI.2016.7451750

[25] R. C. Mayer, J. H. Davis, and F. D. Schoorman. 1995. An integrative model of organizational trust. (1995), 709–734.

[26] B. M. Muir and N. Moray. 1996. Trust in Automation: Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39 (1996), 429–460. https://doi.org/10.1080/00140139608964474

[27] Lin Patrick, Keith Abney, and Ryan Jenkins. 2017. *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence.* Oxford Scholarship Online. https://doi.org/10.1093/oso/9780190652951.001.0001

[28] Teresa Pintado, Joaquin Sanchez, Sonia CarcelÃĺn, and David Alameda. 2017. The Effects of Digital Media Advertising Content on Message Acceptance or Rejection: Brand Trust as a Moderating Factor. *Journal of Internet Commerce* 16, 4 (2017), 364–384. https://doi.org/10.1080/15332861.2017.1396079

[29] Jaime A. Rincon, Angelo Costa, Paulo Novais, Vicente Julian, and Carlos Carrascosa. 2018. A new emotional robot assistant that facilitates human interaction and persuasion. *Knowledge and Information Systems* (28 Jun 2018). https://doi.org/10.1007/s10115-018-1231-9

[30] J. M. Ross. 2008. Moderators of trust and reliance across multiple decision aids (Doctoral dissertation), University of Central Florida, Orlando. (2008).

[31] Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters. 2017. Human Perceptions of the Severity of Domestic Robot Errors. In *International Conference on Social Robotics (ICSR).* Springer, Tsukuba, Japan.

[32] Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters. 2017. 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).* Springer, Tsukuba, Japan.

[33] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joublin. 2013. To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability. *International Journal of Social Robotics* 5, 3 (2013), 313–323.

[34] Sichao Song and Seiji Yamada. 2017. Expressing Emotions Through Color, Sound, and Vibration with an Appearance-Constrained Social Robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17).* ACM, New York, NY, USA, 2–11. https://doi.org/10.1145/2909824.3020239

[35] Dag Sverre Syrdal, Kheng Lee Koay, M. GÃącsi, Michael L. Walters, and Kerstin Dautenhahn. 2010. Video prototyping of dog-inspired non-verbal affective communication for an appearance constrained robot. In *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication.* 632–637. https://doi.org/10.1109/ROMAN.2010.5598693

[36] S. H. Tseng, J. H. Hua, S. P. Ma, and L. e. Fu. 2013. Human awareness based robot performance learning in a social environment. In *2013 IEEE International Conference on Robotics and Automation.* 4291–4296. https://doi.org/10.1109/ICRA.2013.6631184

[37] J. Wainer, D. J. Feil-seifer, D. A. Shell, and M. J. Mataric. 2006. The role of physical embodiment in human-robot interaction. In *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication.* 117–122. https://doi.org/10.1109/ROMAN.2006.314404

[38] Michael L. Walters, Dag S. Syrdal, Kerstin Dautenhahn, René te Boekhorst, and Kheng Lee Koay. 2008. Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots* 24, 2 (01 Feb 2008), 159–178. https://doi.org/10.1007/s10514-007-9058-3

[39] Kun Yu, Shlomo Berkovsky, Ronnie Taib, Dan Conway, Jianlong Zhou, and Fang Chen. 2017. User Trust Dynamics: An Investigation Driven by Differences in System Performance, Vol. 126745. ACM, 307–317.