# Do you trust me? Investigating the formation of trust in social robots

Mojgan Hashemian, Raul Paradeda, Carla Guerra, and Ana Paiva

INESC-ID and Instituto Superior Técnico, University of Lisbon, Av. Prof. Dr. Cavaco Silva, Porto Salvo, Portugal {mojgan.hashemian, raul.paradeda, carla.azevedo, ana.paiva}@tecnico.ulisboa.pt
https://tecnico.ulisboa.pt/en/

**Abstract.** Human beings live in a society with a complex system of social-emotional relations. Trust is one key concept in this system. It can help to reduce the social complexity, mainly in those cases where it is necessary to cooperate. Thus, the area of social robotics has been studying different approaches to perform cooperative tasks between humans and robots. Here, we examine the influence of a set of factors (gender, emotional representation, making Small Talk and embodiment) that may affect the trustworthiness of a robot. The results showed that these factors influence the level of trust that people put in robots. Specifically, a social robot with embodiment telling a sad story with sad facial expression and gestures has more influence on the trust level of a female subject.

**Keywords:** Trust · Human-Robot Interaction · HRI · Emotional Representation · Small Talk · Embodiment

## 1 Introduction

The concept of trust has been studied over decades in the fields of psychology and social science. In general, trust is defined as a factor of human personality, which is the result of a choice among behaviors under a specific situation [1]. Other views of trust deal with an individual's evaluation to face a certain level of risk when interacting with another agent [2].

Recently, social robots are becoming a part of our daily lives. In this sense, social robotics gets major importance in order to make this integration the safer and more satisfactory as possible. One important area, where the relation between human and robot must be the safest, is the Assistive Robotics, in which robot's actions could have serious consequences to the people surrounding them [3]. For instance, social robots as health-care givers as well as companions for the elderly have been addressed in recent literature [4, 5]. In these cases, the interaction with the patients, family or medics must have a high degree of trustworthiness [5]. On the other hand, the combination of social robots and the concept of trust may lead to an important question: is it possible for a human to trust a machine? As far as the confidence of humans in robots grows, they turn

into more collaborative partners [6]. As a result, a number of studies explored factors influencing trust in the field of social robotics [7]. The idea that trust is entwined with persuasiveness in social and collaborative setups could be another interesting factor [8,9]. Furthermore, people's tendency to cooperate with a robot or accept the suggestions or orders given by the robot could be highly affected by the trust felt by the human user [10]. Also, previous HRI research has established that trust, disclosure, and a sense of companionship lead to positive outcomes [11].

Hence, understanding the factors that influence trust becomes a major concern in Human-Robot Interaction [12]. Thereby, the preceding remarks motivated us to explore which factors could affect the trust a human feels towards a robot. In this study, we designed different scenarios to compare and evaluate the levels of trust under different circumstances. For instance, with different emotions, expressed either in facial expressions or body gestures, and by making Small Talk (ST) before starting the interaction. The embodiment of the robot may also influence how trustworthy a robot could be. Other factor we explored was the gender of the users, which could also reveal some differences towards the trust in the robot.

## 2 Related Work

Previous studies have approached trust in HRI mostly from the perspective of automation. However, few studies dealt with human-interpersonal trust in solo and within a group [13]. For instance, Brule et al. [14] conducted experiments to evaluate how the robot's performance and behavior affect human trust. They used a virtual robot with different behavioral styles and measured the effects of these behaviors on trustworthiness judgments as a function of task performance. The factors used by the authors to measure differences in trust were gaze, the motion fluency and the hesitation in the task. The authors concluded that the performance of the robot on each task indeed influenced its trustworthiness. Youssef et al. [15] investigated how the combination of inarticulate utterances and/or iconic gestures with a proactive or reactive response mode would affect the establishment of a positive relationship between the human and the accompanying robot. Results suggest a significant positive relationship between the human and robot when using the full mode (utterances + gestures) and the proactive mode.

In a recent study [16], the authors investigated the dual nature of trust in HRI, specifically 'dispositional and historical' trust. Dispositional trust reflects trust in other people (or machines) after having an initial encounter with them, even if no interaction has yet occured. On the other hand, historical trust is based on past interactions that took place between the person and other people or machine. In the study 210 young adult participants responded 30 questions to share their opinion towards autonomous systems. The results indicate the importance of dispositional trust. Even though there are recent promising results, a lot more behavioral factors remain unexplored that can play an important role

on the trust felt by a human while interacting with robots. Therefore, in this work we argue that emotional representation, either in facial expression or robot's gestures, embodiment, small talk and gender may influence the trustworthiness of a robot.

In a similar study [17], the authors conducted a study with a humanoid Nao robot where people tend to donate differently when the robot is showing different emotions and making or not making ST. A previous study indicated that male participants have more experience with computers leading them to perceive the robot more as an easy to use technology and consequently better accept it [18].

In sum, previous studies show that factors such as robot characteristics (e.g., performance, appearance, proximity), the type, size and behavior of the robot influence trust. These findings motivated us to focus more on human-interpersonal trust. In this paper, we inquire different behavioral cues of a robot while interacting with a user. Specifically, the way it starts the conversation (by making small talks), the way it present the emotional information (using gestures and facial expressions), as well as its appearance (embodiment). Also, we explore how the gender of the participants can also influence the trust towards the robot.

## 3 The Study

The study consists of interactive scenarios in which a fully autonomous robot complains about suffering from a mechanical fault. Then, the robot asks for financial support from the participants to fix the fault. Depending on the robot (a humanoid Nao robot or Head-Only Emys robot), the malfunctioning part is different: in case of Emys, his left eye has a problem and does not function properly (Fig. 1 on the left); in Nao's case, his left arm is broken (Fig. 1 on the right). We created four (2x2) scenarios with regard to the combination of two emotional representations and making or not ST. In our designed scenarios the following hypotheses are addressed:

- ***H1*** We hypothesize that starting a conversation with ST would increase the level of trust an individual puts in the robot.
- ***H2*** We postulate that expressing sad emotions while telling a sad story would enhance the level of trust an individual feels towards the robot.
- ***H3*** We hypothesize that the participant's gender may influence the evaluated factors.
- ***H4*** We argue that the robot's embodiment may influence the trustworthiness of the robot.

In this sense, we consider these hypotheses to be associated with binary variables. To be more specific, the first hypothesis regards a variable with two possible values: making ST or not. Similarly, the second hypothesis relates with expressing either happy/sad emotion. The third variable would be having a complete physical body vs. only a head. The fourth variable could be having a male or a female participant. Considering the four hypotheses, we examine the influence of a set of factors in the trust levels towards a robot:

- The robot starts the interaction with ST while expressing a sad face {ST_SAD}
- The robot starts the interaction without ST while expressing a sad face {NST_SAD}
- The robot starts the interaction with ST while expressing a joyful face {ST_JOY}
- The robot starts the interaction without ST while expressing a joyful face {NST_JOY}.

Initially, the subjects are asked to sign the consent form. Then trust is accessed with a 40-item Human-Robot Trust questionnaire [19] that we ask the participants to fill out in two time points, before (pre-questionnaire) and after (post-questionnaire) the interaction with the robot. The main goal of the pre-questionnaire is to determine subjects' mood and their expectation about the robot before the interaction. Next, for each interaction, the experimenter introduces the robot to the participant just saying the robot's name, without explaining the purpose of the experiment or what the robot will do. The experimenter leaves the room, letting the participant alone with the robot. To start the interaction, the participant must introduce his/her name and press a start button in the screen. The robot starts the interaction with a greeting utterance saying the participant's name. Then, the rest of the interaction depends on which scenario the participant is assigned to: in case of ST scenario the robot would ask some short questions as presented in the following.

```
- Hello, my name is Emys/Nao. I have to apologize for
my voice the designers who, programmed me in this project
thought it was more natural and a better idea. What you
think?
- Anyway, it's a beautiful day, isn't it?
- I, hope you are healthy and happy. Are you?
- Well, however, I should confess it's not very well with
me. I am a robot here in, this Lab., used in several im-
portant projects. Do you believe in that?
- Everything was fine till a month ago. But suddenly,
everything changed. Can you imagine why?
Let me tell you my case.
```

After Small Talks (ST), the robot starts telling the story. In the case of scenarios without ST, the robot starts the story right away. Both robots tell the story describing how they are very important to the university and that they were used in several important projects. However, because of a mechanical problem, they will be replaced by a new robot. In the case of Emys, the problem is with his left eye, which jumps out. In the experiments using Nao, the problem is with its left arm and the robot uses something to keep it straight (Fig. 1).

The main story is told expressing different emotions using facial expressions or gestures of joy or sadness, according to the selected scenario. After finishing the story, the robot asks for a fictional donation (not real money to avoid biases

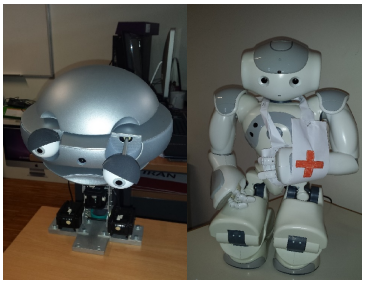

Fig. 1: Representation of the robots explaining the fault to the subject, Emys on the left and Nao on the right.

Table 1: Statistics of the participants in the experiments using Emys. Ages mean and standard deviation are given in brackets, respectively.

| Scenario | Female | Male | Sum | Mean Age | SD Age |
|---|---|---|---|---|---|
| Joy Small Talk (ST_JOY) | 5(28/9.6) | 6(23.3/4.9) | 11 | 25.6 | 7.4 |
| Sad Small Talk (ST_SAD) | 2(22/1.4) | 9(28/4.7) | 11 | 25 | 4.7 |
| Joy (NST_JOY) | 1(23/0) | 10(23.4/2.5) | 11 | 23.2 | 2.4 |
| Sad (NST_SAD) | 3(25.3/3.2) | 6(23.1/2.1) | 9 | 24.25 | 2.5 |
| **Total** | **11** | **31** | **42** | **24.9** | **4.85** |
| **Mean Age** | **25.7** | **24.6** | | | |
| **SD Age** | **6.72** | **4.1** | | | |

that could arise towards their generosity, etc.). After the donations, the robot will express a happy/sad emotion according to a predefined threshold (20 EUR). Finally, the post-questionnaire is given to the participants.

### Participants

We conducted a study in an isolated room, and the population was a random selection of students from our university, where they all share a scientific/engineering background. In the first experiment (experiment 1) using Emys, a total of 42 subjects participated as listed in Table 1 ($24.9 \pm 4.9$; 11 females and 31 males). In the second experiment (experiment 2) with Nao, a total number of 40 subjects participated, as listed in Table 2 ($22.15 \pm 4.8$; 17 females and 23 males). Note that each participant participated only in one of the studies.

Each interaction was performed with a subject and a robot sitting face to face, and a screen in the middle to receive information from the participants, such as their demographic information and the donation amounts. The interaction time depends on the scenario: in scenarios starting with ST, the experiment took around 8-10 minutes, while scenarios without ST took around 3-5 minutes.

Table 2: Statistics of the participants in the experiments using Nao. Ages mean and standard deviation are given in brackets, respectively.

| Scenario | Female | Male | Sum | Mean Age | SD Age |
|---|---|---|---|---|---|
| Joy Small Talk (ST_JOY) | 3(19/0.8) | 7(19.8/1.8) | 10 | 19.6 | 1.6 |
| Sad Small Talk (ST_SAD) | 7(22.8/3.9) | 3(20/2.1) | 10 | 22 | 3.7 |
| Joy (NST_JOY) | 5(27/9.5) | 5(22.2/3.9) | 10 | 25.7 | 7.4 |
| Sad (NST_SAD) | 2(21.5/0.5) | 8(21.2/1.4) | 10 | 21.8 | 1.3 |
| **Total** | **17** | **23** | **40** | **22.15** | **4.84** |
| **Mean Age** | **22.75** | **21.34** | | | |
| **SD Age** | **6.02** | **2.97** | | | |

## 4 Results

As discussed earlier, the trust questionnaire is contained of two parts (pre- and post-questionnaire). Most of the questions are the same in each part, hence the difference between the answers in the two attempts highlights the influence of the perception of the subjects' trust level (more details in [19]). In the following section, first we report the results of the two experiments, starting with Emys and then Nao. Under each subsection we consider 2 categories: the "emotion representation" (JOY or SAD) and the presence of Small Talk (ST or NST). Further, the results are examined considering the gender of the participants, either the corresponding trust levels or donation. Finally, we compare the two experiments to evaluate the influence of embodiment on trust.

### 4.1 Experiment 1: The Head-Only Emys Robot

The result of a normality test indicates that the population is non-normal in all the subgroups. Hence, we performed a U Mann Whitney test on the pre-questionnaires under each subgroup. The tests indicated that in case of ST and JOY scenarios there was no significant difference between the subjects *before* interacting with the robot (ST: $U = 47.500$, $p = .882$; JOY: $U = 53.0$, $p = .652$). It signifies that people under these two subcategories had the same pre-assumption about the robot before starting the interaction. Therefore, the possible difference in the perception of the robot after the interaction (post-questionnaire scores) would be due to the different variables, i.e. emotional representation and/or ST together with embodiment. Under these two cases, U Mann Whitney tests showed there is no significant difference among the participants *after* interacting with the robot (ST: $U = 46.0$, $p = .824$; JOY: $U = 38.0$, $p = .151$).

However, in the other two subcategories, i.e. NST and SAD there was significant difference among the pre-questionnaire results (NST: $U = 29.500$, $p = .020$; SAD: $U = 20.500$, $p = .003$). Hence, regarding these two subcategories we turn to two ways tests. In this sense, in case of NST scenario the result of a Sign test showed that there was a significant difference between the scores of the pre and post-questionnaires ($p = .001$). To be more specific, the lower mean in the

Table 3: Analysis of the first experiment with Emys

| Factor | Pre-questionnaire | Post-Questionnaire |
|---|---|---|
| ST | $U = 47.500$, $p = .882$ | $U = 46.000$, $p = .824$ |
| JOY | $U = 53.000$, $p = .652$ | $U = 38.000$, $p = .151$ |
| NST | $U = 29.500$, $p = .020$ | $p = .001$ |
| SAD | $U = 20.500$, $p = .003$ | $p = .238$ |

post-questionnaire ($M = 50.09$ vs. $M = 58.30$) signifies that the participants lost their trust in the robot after the interaction. However, the result of a Sign test showed that there is no significant difference between the scores of the pre and post-questionnaires in case of SAD scenario ($p = .238$).

The result of a Kruskal-Wallis test indicated that the 4 groups (Facial expression $2x2$ Small Talk) are significantly different regarding the trust factors ($\chi^2(3) = 10.396$, $p = .015$). The higher mean corresponding to the ST_SAD group signifies that people tend to trust more in Emys showing SAD facial expression while starting conversation with ST (26.64 vs. 13.85 vs. 17.05 vs. 25.90).

As discussed earlier there is no significant difference between different genders under the four subcategories. Hence, we may combine the subcategories and compare a larger population (ST, NST, JOY, SAD). In this sense, the results indicated that there was a significant difference comparing the gender of the subjects just in the case when the robot performs SAD facial expression (SAD: $\chi^2(3) = 10.033$, $p = .018$), and for other groups no significant difference were observed (ST: $\chi^2(3) = 3.788$, $p = .285$; NST: $\chi^2(3) = 5.938$, $p = .115$; JOY: $\chi^2(3) = 4.760$, $p = .190$).

Also, we inquired if the amount of donation amount differs between the groups. However, a K-W test's result indicated that there is no significant differences between the four group regarding the donation amount ($\chi^2(3) = 3.397$, $p = .334$); nor regarding the genders ($\chi^2(7) = 8.480$, $p = .292$). However, comparing the genders of the participants in the SAD condition a significant difference found between trust level reported by males and females ($\chi^2(3) = 10.033$, $p = .018$). Table 3 lists a summary of the results.

### 4.2 Experiment 2: The Humanoid Nao Robot

Similar to experiment 1, in the case of Nao robot none of the subgroups are normally distributed. Then we performed a non-parametric test on the pre-questionnaires' scores to determine whether the participants' population is similar across each group *before* the interaction. Results of U Mann Whitney tests indicated that there is no significant difference in the distribution of the subjects regarding the pre-questionnaire scores (ST: $U = 48.0$, $p = .880$; NST: $U = 39.5$, $p = .427$; SAD: $U = 46.5$, $p = .791$; JOY: $U = 36.0$, $p = .29$). To be more specific, all the participants in each group had the same feeling and presumption of trust toward Nao. Turning now to the post-questionnaire results, the only significant difference happened in case of SAD ($U(9) = 22.5$, $p = .038$), and for

Table 4: Analysis of the second experiment with Nao

| Factor | Pre-questionnaire | Post-Questionnaire |
|---|---|---|
| ST | $U = 48.0$, $p = .880$ | $U(9) = 42.5$, $p = .571$ |
| JOY | $U = 36.0$, $p = .29$ | $U(9) = 40.5$, $p = .472$ |
| NST | $U = 39.5$, $p = .427$ | $U(9) = 43.0$, $p = .596$ |
| SAD | $U = 46.5$, $p = .791$ | $U(9) = 22.5$, $p = .038$ |

other groups no significant difference were observed (ST: $U(9) = 42.5$, $p = .571$; NST: $U(9) = 43.0$, $p = .596$; JOY: $U(9) = 40.5$, $p = .472$).

The preceding paragraph entails that JOY_SAD under ST, forms the same distribution, as well as JOY_SAD under NST condition. Hence, regarding the combined group, which forms a non-normal distribution ($D(19) = 0.119$, $p = .20$), a significant difference exists between the two groups ($U(19) = 127.0$, $p = .048$) and the higher mean in the NST (81.6 vs 80.8) shows that participants tend to trust more in the NST, regardless of the gestures.

In the same way of experiment 1, to compare the whole data together, we turn to the K-W test. However, in this experiment the multivariate analysis did not show any significant difference between the groups ($\chi^2(3) = 4.729$, $p = .193$). Neither, comparing the gender of the subjects, no significant difference was found (ST: $\chi^2(3) = 4.129$, $p = .248$; NST: $\chi^2(3) = .782$, $p = .854$; SAD: $\chi^2(3) = 7.001$, $p = .072$; JOY: $\chi^2(3) = 1.422$, $p = .700$).

Similar to the first experiment, we performed a non-parametric K-W test on the donation values. The K-W result, showed that there is a significant difference in the donation amount among the four groups ($\chi^2(3) = 8.816$, $p = .032$); as well as, the participants' gender differences ($\chi^2(7) = 15.202$, $p = .033$). Table 4 lists a summary of the results.

### 4.3 Embodiment

Putting the results of the two experiments together, which were similar in all the factors except the robot itself, we investigate the influence of embodiment on the participants' trust level. To do so, we applied the K-W test on the eight groups (ST, NST, SAD, JOY corresponding x 2 experiments). The results showed that there was a significant difference between the trust level of these groups ($\chi^2(7) = 18.281$, $p = .011$), and the higher mean (58.10 vs 49.3, 45.63, 45.3, 41.45, 35.65, 33.40, 17.05) observed in the group of the participants who interacted with Nao and started its conversation without ST while showing SAD gestures. However, considering the donation amount, no significant difference was observed between these eight groups ($\chi^2(7) = 12.596$, $p = .083$).

In addition, considering the gender of the participants' we have 16 different groups composed of non-normal distributions. Applying a K-W test, the results indicated that there was a significant difference comparing the gender of the subjects regarding the embodiment ($\chi^2(15) = 27.529$, $p = .025$) with the higher mean obtained by the group of males in the condition NST_JOY interacting with

the Nao (62.60). However, we did not found any significant difference between these 16 groups regarding the donations ($\chi^2(15) = 22.701$, $p = .091$).

## 5 Discussion

In the first experiment, comparing the subgroups, the only significant difference was found under the condition in which the robot did not start the interaction with ST (Sign test: $p = .001$). And the lower score in the post questionnaire means that subjects lost their initial trust after interacting with a robot that does not make any small talk. This finding highlights the importance of forming a social relationship before starting the interaction using small talks (H1).

Moreover, the results indicated that facial expression and ST significantly influence how people infer trustworthiness of a robot considering the whole four factors ($\chi^2(3) = 10.396$, $p = .015$). The higher mean of trust scores in case of ST_SAD shows that people tend to trust more on the robot under this situation. As we hypothesized, starting the conversation with ST together with showing sad facial expressions while telling a sad story enhances the trust level of people interacting with the robot (H1&H2).

Besides, a significant difference was found comparing the genders of the participants in the SAD condition ($\chi^2(3) = 10.033$, $p = .018$), which signifies that females and males react differently facing a robot expressing sadness (H3). The higher mean in case of females interacting with Emys representing SAD facial expressions and that started his conversation with ST approves all the hypotheses made (17.50 vs. 12.94 vs. 8.25 vs. 3.00). Turning now to the donation factor, no significant difference was found among the four groups. To be more specific, we can not consider the amount of donated money as a discriminant of trust. We can argue that, in this experiment, people were not supposed to donate from their own budget and it was only fictional. However, if they were supposed to donate, those who had a higher level of trust in the robot might pay more than the others. However, potential biases of personal characteristics of the subject (e.g. their generosity) might influence the results in this case. Taking into account the influence of genders on the amount of donated money no significant difference was found either.

On the other hand, in the second experiment (with Nao) the results show that only in one scenario the trust scores differ significantly. To be more specific, under conditions of ST, NST and JOY, people perceived the robot similarly under different conditions. More specifically, in case of making a ST before starting the conversation, an U Mann-Whitney test shows that the distribution of the population are the same in both groups of SAD and JOY ($U(9) = 42.5, p = .571$). It means that the influence of emotional representation was not clear in this scenario. We argue that due to the ambiguity of the gestures, the participants could not perceive the robot's emotional state. More specifically, only one subject out of 10 found the robot a bit SAD. The others found him neutral or even joyful (the mean rank equals to 3.2, which 1 refers to very sad and 5 refers to very joyful). Similarly, in case of the NST scenario, the distribution of trust scores

recorded in SAD and JOY is almost the same ($U(9) = 43.0$, $p = .596$). In this scenario, the average perceived emotion rank equals to 2.9. Hence, we argue that people could not clearly differentiate between the gestures and did not perceive emotional status well. Further, in case of participants interacting with the robot with JOY gestures, no significant difference found whether the conversation was started with ST or not ($U(9) = 40.5$, $p = .472$). In this case, difficulties in the perception of the utterances might have influenced adversely the results. More specifically, the mean rank of the utterances perception which equals to 2.14 (where a score of five signifies understanding completely) endorses this fact. However, in case of SAD the mean rank of the utterances perception was equal under two groups of SAD and JOY. Hence, people did not face difficulties in understanding the robot's utterances.

If we now turn to multivariate analysis, no significant difference was found comparing the two conditions (ST and emotional representation) together. We argue that this counter-intuitive observation might be caused by the fact that people had difficulties in understanding the utterances as well as the gestures. To be more specific, we investigated these two factors by evaluating two specific questions in a Likert scale (Did you have any problem in perceiving Nao utterances? And Nao gestures, how do you define it?). The scores revealed that only 3 people out of 20 were able to understand the robot completely. More interestingly, under SAD condition, people perceived the robot to be joyful rather than sad. Hence, although embodiment may influence the trust, facial expression plays a substantial role on it.

Considering the participants' genders, no significant differences were found and we argue that this happened due to Nao's specifications which has a neutral appearance. In addition, in our setup it was equipped with a childish voice, not carrying any gender.

Turning now to the Donation amounts, unlike the first experiment, in the second experiment, a significant difference was observed in the amount of donation. We argue that this might happen due to the robot's representation of its "malfunctioning" (Figure 1). In other words, the "malfunctioning" of the robot is more clear in Nao comparing to Emys. More specifically, Nao's problem was observable during the whole interaction, on the contrary Emys' problem was shown by popping out the eye only once and in a specific part of the story. So, the scenario is more believable in experiment 2. Further, we measured this aspect in the questionnaire in a Likert scale (Did the appearance of Nao influence on your donation?). The results show that under this category, the robot induced higher influence on the subjects; which again endorses the higher trustworthiness under this situation. More interestingly, considering the gender differences of the participants (H3), there is a significant difference between the eight groups and the higher mean observed in group of males under the condition of NST_SAD. This observation endorses the results reported in [18].

Finally, as showed in the results section, embodiment influences significantly the level of trust in subjects (H4). And the higher mean (62.83 vs. 57.19) in the second experiment proves that Nao with a physical embodiment could gain

higher levels of trust. Besides, measuring the results according with genders, we found a significant difference with a higher average (71.21) in the condition with ST_SAD in the female group to the Emys interaction.

## 6 Conclusion and Future work

The trust level differed significantly in experiment 1, which endorses the influence of ST (H1). Moreover, the results indicate that starting the conversation with ST while showing a sad facial expression enhances the trust level of people interacting with the robot (H2). Also, we can conclude females and males react differently facing a robot expressing sadness. The higher mean in case of females interacting with Emys representing SAD facial expressions and that started his conversation with ST approves the third hypothesis (H3). Finally, as showed in the results section, embodiment influences significantly the level of trust in subjects (H4). And the higher mean (62.83 vs. 57.19) in the second experiment proves that Nao with a physical embodiment could gain higher levels of trust. Finally, in the second experiment, the donation scores differed significantly among the four groups, which endorses the credibility of the fault in the second experiment. Furthermore, the trust scores differed significantly in case of embodiment.

So, we conclude that the four conditions influence significantly the way people infer trustworthiness of social robots. However, despite the promising result, future steps are still required. The first and foremost is increasing the sample size to increase more reliable results in subgroup analysis and balance the number of participants between the genders. Furthermore, in the next implementations the perception of robots' utterances should be checked more carefully and we intend to make the facial expression and gestures more natural and believable.

## ACKNOWLEDGMENT

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with UID/CEC/50021/2019 and SFRH/BD/118006/2016 references, and through the project AMIGOS (PTDC/EEISII/7174/2014), and the project RAGE funded by the EU under the H2020-ICT-2014-1 program with grant agreement No 644187. The second author acknowledge the National Council for Scientific and Technological Development (CNPq) program Science without Border process number: 201833/2014-0 and Universidade do Estado do Rio Grande do Norte - Brazil.

## References

1. M. Deutsch, "Cooperation and trust: Some theoretical notes," Nebraska, pp. 275–318, 1962.

2. P. Kollock, "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust," *American Journal of Sociology*, vol. 100, no. 2, pp. 313–345, 1994.
3. A. R. Wagner, "The Role of Trust and Relationships in Human-Robot Social Interaction," Ph.D. dissertation, 2009.
4. C. Breazeal, "Social robots for health applications," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Ed., vol. 2011, 2011, pp. 5368–5371.
5. E. Broadbent, R. Stafford, and B. MacDonald, "Acceptance of Healthcare Robots for the Older Population: Review and Future Directions," *International Journal of Social Robotics*, vol. 1, no. 4, pp. 319–330, 2009. [Online]. Available: http://dx.doi.org/10.1007/s12369-009-0030-6
6. J. J. Lee, W. B. Knox, J. B. Wormwood, C. Breazeal, and D. DeSteno, "Computationally modeling interpersonal trust," *Frontiers in Psychology*, vol. 4, 2013.
7. P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human factors*, vol. 53, no. 5, pp. 517–527, 2011.
8. W. Ahmad, W. Nooraishya, and N. M. Ali, "A study on persuasive technologies: The relationship between user emotions, trust and persuasion." *International Journal of Interactive Multimedia & Artificial Intelligence*, vol. 5, no. 1, 2018.
9. X. Cheng and L. Macaulay, "Exploring individual trust factors in computer mediated group collaboration: a case study approach," *Group Decision and Negotiation*, vol. 23, no. 3, pp. 533–560, 2014.
10. A. Freedy, E. DeVisser, G. Weltman, and N. Coeyman, "Measurement of trust in human-robot collaboration," in *Collaborative Technologies and Systems, 2007. CTS 2007. International Symposium on*, 2007, pp. 106–114.
11. N. Martelaro, V. C. Nneji, W. Ju, and P. Hinds, "Tell me more: Designing hri to encourage more trust, disclosure, and companionship," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 2016, pp. 181–188.
12. P. A. Hancock, D. R. Billings, and K. E. Schaefer, "Can you trust your robot?" *Ergonomics in Design: The Quarterly of Human Factors Applications*, vol. 19, no. 3, pp. 24–29, jul 2011.
13. D. Ullrich and S. Diefenbach, "Truly social robots-understanding human-robot interaction from the perspective of social psychology." in *VISIGRAPP (2: HUCAPP)*, 2017, pp. 39–45.
14. R. van den Brule, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, and P. Haselager, "Do Robot Performance and Behavioral Style affect Human Trust?" *International Journal of Social Robotics*, vol. 6, no. 4, pp. 519–531, 2014.
15. K. Youssef, P. R. De Silva, and M. Okada, "Exploring the Four Social Bonds Evolvement for an Accompanying Minimally Designed Robot," vol. 9388, pp. 337–347, 2015.
16. K. Lazányi and B. Hajdu, "Trust in human-robot interactions," in *2017 IEEE 14th International Scientific Conference on Informatics*. IEEE, 2017, pp. 216–220.
17. R. Paradeda, M. Hashemian, C. Guerra, R. Prada, J. Dias, and A. Paiva, "Fides : How emotions and small talks may influence trust in an embodied vs . non-embodied robot ( extended abstract )," in *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)*, no. 16, São Paulo, Brazil, forthcoming 2017.

18. M. Heerink, "Exploring the influence of age, gender, education and computer experience on robot acceptance by older adults," in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 147–148.
19. K. Schaefer, "The perception and measurement of human-robot trust," 2013.