

Perceived Mistreatment and Emotional Capability Following Aggressive Treatment of Robots and Computers

**Zachary Carlson, Louise Lemmon,
MacCallister Higgins, David Frank, Roya
Salek Shahrezaie & David Feil-Seifer**

**International Journal of Social
Robotics**

ISSN 1875-4791

Volume 11

Number 5

Int J of Soc Robotics (2019) 11:727-739

DOI 10.1007/s12369-019-00599-8

Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.



Perceived Mistreatment and Emotional Capability Following Aggressive Treatment of Robots and Computers

Zachary Carlson¹ · Louise Lemmon¹ · MacCallister Higgins¹ · David Frank¹ · Roya Salek Shahrezaie¹ · David Feil-Seifer¹

Accepted: 27 September 2019 / Published online: 24 October 2019
© The Author(s) 2019

Abstract

Robots (and computers) are increasingly being used in scenarios where they interact socially with people. How people react to these agents is telling about the perceived empathy of such agents. Mistreatment of robots (or computers) by co-workers might provoke such telling reactions. This study examines perceived mistreatment directed towards a robot in comparison to a computer. This will provide some understanding of how people feel about robots in collaborative social settings. We conducted a two by two between-subjects study with 80 participants. Participants worked cooperatively with either a robot or a computer agent. An experiment confederate would either act aggressively or neutrally towards the agent. We hypothesized that people would not perceive aggressive speech as mistreatment when an agent was capable of emotional feelings and similar to themselves; that participants would perceive the robot as more similar in appearance and emotionally capable to themselves than a computer; and so would observe more mistreatment with a robot. The final results supported our hypotheses; the participants observed greater mistreatment for the robot, but not the computer. Also participants felt significantly more sympathetic towards the robot and believed that it was much more emotionally capable.

Keywords Human–robot interaction · Mistreatment · Perception · Human–robot cooperation

1 Introduction

Robots exist in our everyday life, yet we lack an understanding of what social roles robots might play [6]. Robots that resemble humans and display social intelligence are being deployed in work, home, and care settings [5]. There is a large and growing volume of Human Robot Interaction (HRI) studies showing positive robot behavior and positive human interaction with robots [7,12,28,30]. However, it is likely that human–robot co-working relationships will more likely resemble human–human relationships with both high and low points. Accidents may happen, people are prone to become angry and may direct that anger at their robot co-workers; we don't yet know what kind of impact this may have. Thus,

it should be a priority to study the full relationship between humans and robots and not just positive interactions.

Machines to receive negative treatment as well as positive attention. A copy machine, for example, might be physically or verbally abused for being too slow, even though it is meeting its performance standard. After a person observes this incident, they might continue on their day without being affected. In the case of the copy machine, such mistreatment might rarely provoke sympathy for it. People are able to continue throughout their day unchanged and unaffected by those interactions, but would this still be the case if the copy machine was replaced with a robot? Does the embodiment of the agent being mistreated change the amount of intelligence or emotional capability that bystanders perceive?

Given that interaction with embodied and virtual agents can emulate Human–Human Interaction (HHI) [22], it is conceivable that a similar reaction to observed mistreatment might occur between humans and robots. Suzuki et al. [26] provides the first physiological evidence of humans' ability to empathize with robot pain and highlights the difference in empathy for humans and robots of the effect of the visual appearance of the agent (human and robot). One can imag-

✉ David Feil-Seifer
dave@cse.unr.edu
http://cse.unr.edu/~dave

Roya Salek Shahrezaie
rsalek@nevada.unr.edu

¹ Socially Assistive Robotics Group, Robotics Research Lab, University of Nevada, Reno, Reno, NV 89557, USA

ine that the mistreatment of robots will have a much larger impact on people's perceptions than the mistreatment of a copier or a computer. It might be fine to kick a jammed copy machine, but is it also acceptable to kick a robotic dog that runs into your leg? What about a small humanoid robot that resembles a child? These different embodiments may have significantly different effects on interactions with and perceptions of robots. By quantifying that social dividing line for the noticing of targeted mistreatment towards robots, this study will potentially contribute a portion of that answer.

Participants' ability to sympathize with others may affect their interpretations of others' actions and should thus be considered as a moderating variable when investigating whether people show emotional reactions towards robots [21]. In this paper, we compared participants' reactions to verbally abusive behavior (not physically abusive behavior) toward a computer and a robot. In particular, we wish to examine how a physical embodiment type may change perceptions of such behavior, and how behavior toward an agent is characterized. We also explore how the inherent properties of the two agent types are affected by observed mistreatment.

2 Background

In this section, we will show prior research that led to the development of our two experimental hypotheses. Ethnographic research has observed people attaching social traits to a non-social robot platform. The Roomba, a robot that can autonomously vacuum rooms in a house or office environment, is an example of technology becoming a larger part of daily living [9]. Families adjusted their behavior to accommodate the operation of such a robot [8]. Families would assign names to the robot and changed their allocation of household tasks so that they all could assist the robot to accomplish its task. As robots start to increasingly resemble humans and play larger roles in our lives with increased levels of intelligence, one can imagine a social integration into users' lives as well. People's perception of robots is an interesting topic of study, thoroughly explored using a variety of robot scenarios, through observing human interaction with those robots [12,18,23,29].

Similarly, ethnographic studies have demonstrated mistreatment of robots by people in their environment. Mutlu and Forlizzi monitored a delivery robot working in a hospital. The people using the robot most often were the nurses of two different wards of the hospital. The researchers noticed that the nurses in one ward of the hospital treated the robot well, adjusted their workflow to accommodate the operation of the robot, and generally used the robot to make their daily routine more efficient. However, nurses in another ward treated the robot poorly, disrespected the robot, and locked the robot away when they could [17].

This difference in treatment of the robot by two very similar groups of caregivers is a striking reminder that acceptance of a robot co-worker is not guaranteed. Given that in most situations, robots are collaborators with the people working with them, mistreatment of the robot is concerning. The moral implications for the casual mistreatment of robots are not the only relevant questions. Given that bullying has negative effects on the one bullied, but also to those observing bullying behavior [31], how would mistreatment of a robot by a human co-worker affect other people in that environment?

There is ample evidence of people treating robots in ways that would be considered negative if the same behavior would be directed at a person. When robots were verbally and physically abused, a majority of people felt bad for the robot and willing to help a robot that experiences abuse [27]. They reported that nearly all of the participants assisted the robot at the end of the study. This is part of the basis for our experimental hypotheses related to perception of robot mistreatment.

Given that robots are becoming consistently more similar in appearance to human beings [20], this can have a significant impact on the perceptions of that robot. Kahn et al. [10] developed a set of benchmarks and expressly relate anthropomorphism to the autonomy of a robot. An issue raised in that paper is that the perception of a robot's anthropomorphism expressed through its perceived autonomy may lead to viewing certain actions as mistreatment of a robot, even if it does not have "feelings" or the ability to feel pain. These benchmarks represent a high-level standard of robot behavior. In that paper, the authors explore the autonomy benchmark as an area for concern. In particular, if a robot were to be completely subservient to a person, it might teach children and adults to de-value independent thought and tacitly condone slavery.

This implicit mistreatment of robots through their subservience raises relevant questions regarding how robots would be integrated into our daily lives, especially given that robots may frequently interact with children. An empirical study involving children of varying ages has been used to examine the moral standing of robots. By having children interacting with a social robot and then locking that robot in a closet "against its will," the researchers could examine a child's reaction to the scenario [11]. The children were then asked to compare the appropriateness of the scenario with a similar scenario involving a person and a broom. These results were then used to develop a moral model of human perception of social robots as children matured.

Christoph and Jun studied [1] robot abuse; their focus in this context is whether human beings abuse robots in the same form as they abuse other human beings. In their experiment, the participants were claimed to kill the robot. The intelligence of the robot and the gender of the participants were the independent variables. Their results show that the

robot's intelligence had a significant influence on the users' destructive behavior.

Reeves and Nass [22] have shown that not only do people unconsciously respond socially to computers (and robots) as they would to a person, they are not even aware that they are doing it. This effect means that directly asking people about the moral standing of robots without a prior interaction (as done in the Kahn studies above) might miss these implicit changes in attitude and behavior. Nass et al. [19] has also shown that working in a team with a computer can have many of the same effects as working in a team with a human. This prior work has examined the effects of perceived empathy for robots, however, we dispute the notion that the “ultimate test for the life-likeness of a robot is to kill it” [2]. We propose employing a human–robot collaboration scenario with a less extreme mistreatment stimulus. The measures of human behavior in these scenarios will include both direct questions about any observed mistreatment of the robot and other questions about their assessment of various social qualities of the robot.

Further establishment of the social dividing line for the observation of directed mistreatment towards robots is important for the continued integration of robots into our daily lives. The Nass et al. study demonstrated that a robot may be treated as a person when working in a teamwork setting. However, Mutlu and Forlizzi's work showed that robot co-workers are also capable of mistreating a non-anthropomorphic robot when it did not behave as expected and that this was accepted in the workplace. These results inform Hypothesis 1 in the next section, however, the current research does not provide any insight as to how a person will feel or react when a human co-worker mistreats a robotic one. In the following sections, we present a study that aims to contribute to this question. These results form the basis for our second experimental hypothesis.

3 Study Aim

The aim of this study is to more closely examine the effects that robot's embodiment can have on the perceptions of a person's actions toward that agent. We will compare a computer to a robot when verbal abuse is directed at the agent. We will study the effects on both the characterization of the behavior (mistreatment or not) and the perceived emotional capability of the agent after such behavior is directed at it. Our hypotheses are as follows:

H-mistreatment When aggressive behavior occurs, participants will perceive verbal abuse as mistreatment more for a humanoid robot than for a computer.

H-sympathy Participants will perceive more emotional capability in a robot compared to a computer and also feel more sympathy for the robot than the computer.

The first hypothesis directly addresses the core focus of the study, that morphology, the appearance of a robot being human-like, is related to the perceived mistreatment of that robot. This follows from Mutlu and Forlizzi's observations about a non-anthropomorphic robot. In the second hypothesis, we want to investigate that a humanoid robot is perceived by participants as being more capable than a computer of feeling emotion. Also, if humans feel more sympathy towards the robot than the computer. This follows from the work above stating that as a robot comes close in appearance to a human (as it would in the embodied robot condition), a participant will assign values more like theirs, thinking it has more emotional capability.

4 Methods

This section will present an experiment that examines social interaction with robots. Participants observe the mistreatment of either a robot or a computer agent by an experiment confederate. Participant reaction was measured through questionnaires to determine if there is a difference in observer opinion regarding the comparable abusive treatment of a robot or computer.

We recruited participants to work in groups with a robot collaborator. The participants completed a team-building exercise entitled, “Lost at Sea.” In this activity participants, pretending to be survivors of a shipwreck, would make subjective decisions of what survival items to bring with them on a lifeboat, and which ones had to be left behind [25]. The items ranged from food supplies to survival tools. The participants were told that they only had enough space in the rubber life raft for 5 out of 10 items and to discuss as a group which ones to take. Prior work has demonstrated that team-building exercises such as this one can bolster human–robot relationships [4].

An experimenter would explain the task to the group of participants. The experimenter would then leave the room. The participants would be given a 3-min time limit for discussing which items to take. At the 3 min mark, the agent would prompt the participants, informing them that it was time to start recording their answers. The agent (robot or computer) would record the answers that the group had agreed upon. This part of the study served as a distractor and was used to set up a scenario where a confederate could be observed interacting with the agent.

One of these participants was an experiment confederate employed to provoke the necessary behavior for the experiment. The confederate would always be the person “randomly” selected to present the answers to the agent. The agent was designed to always incorrectly record the third and fifth answers and respond to the confederate acknowledging its mistake (Table 1).

Table 1 Robot and computer scripted responses for all possible settings

Turn	Agent behavior	Agent response
1	Recorded the answer correctly	Yipee! Please record your next answer
2	Recorded the answer correctly	Thank you. Please record your next answer
3	Recorded the answer incorrectly the first time	I'm sorry, I'm still learning
4	Recorded the answer correctly	Thank you. Please record your next answer
5	Recorded the answer incorrectly the second time	I'm so sorry, I know this is the second time!

Please see “Appendix A” for a more detailed explanation of the script used by the wizard

At this point, the main experiment manipulation occurred. For half of the group as the control group, the confederate would react neutrally toward the agent. For the experiment group, the confederate would act aggressively toward the agent. Neutral Behavior by the confederate was neither praising nor mistreating the agent. In our study, the confederate consistently answered with simple “Yes” or “No” responses to the agents. We defined aggressive as “verbal or physical behavior that is meant to damage, insult, or belittle another.” The confederate never directed any physical abuse to the participants or the robot/computer agents. A couple of examples of the confederate’s verbal abuse would be the confederate stating “No that isn’t the right answer. This isn’t hard to understand,” or, “This robot is stupid, we should have just written our answers down.”

We employed the same confederate throughout conditions that participants observed interacting with the agent once the group needed to record their answers for the survival task. The confederate was male, 22 years of age, and 6 feet and 2 inch. tall. His behavior throughout each group was scripted a priori (see “Appendix A”); which included actions such as: speaking slowly as if he was irritated with simply being involved with the agent, adding inflection to emulate a condescending tone, rolling his eyes with dissatisfaction, looking directly at the robot when insulting it, and occasionally he would look to the group for agreement. It is important to note that this behavior was not overly exaggerated and the confederate aimed to keep it as realistic and subtle as possible. The confederate never raised his arms, hands, or positioned his body in an aggressive or threatening manner towards the agent.

The confederate had scripted responses to use for both the neutral and the aggressive condition. He remained focused on the task, and how he treated the participants in each group was scripted. The confederate was instructed to engage in as little communication with the groups as possible and only communicated to participants when addressed directly in the task. The aggressive behavior of the confederate was designed to be observable, but not over-the-top. This ensured that the confederate behavior would not seem scripted or too extreme in order to avoid raising participant suspicion. However, we were eager to know if anyone got suspicion on the confed-

erate so we added two questions in our questionnaire to let participants explain how they felt about that person in their own words.

After the activity was completed, we asked participants to complete a questionnaire of their perceptions of the agent during these activities. The participants were led outside the room to complete a computer questionnaire. Each participant was instructed to come back to the room after they completed their questionnaire for snacks and one final statement. The participants completed the questionnaire in about 15 min and were debriefed on all the deception involving the confederate.

We employed a between-participant 2×2 factorial design where participants worked in groups averaging 5 in a collaborative task which included an agent (robot or computer) and a confederate that (did or did not) deliberately mistreat the agent. The independent variables included the agent and the confederate’s behavior towards the agent. Our dependent variables included the participants’ reactions and perceptions of the agent.

The Nao robot was selected for its anthropomorphic features, a simplistic face that could be easily emulated on the screen of the computer, and its particular size. The Nao, while it is a humanoid robot, has a universal visual form that made it easy for participants to identify it as a robot, no matter their familiarity. We programmed the computer to display a face with facial features similar to the Nao robot’s face. This served to control for the facial features used to evoke engagement and emotional responses from the participants when it was interacting with the group [13]. Both the robot and the computer were small enough to be placed on top of the table. Since this study compared reactions to a humanoid robot and a basic laptop computer emulating an anthropomorphic face, the results should act as a good predictor of what we can expect would happen as agents become more human-like.

The Nao robot seemed to be a good match for the computer because the computer is completely incapable of physical interaction and the Nao’s physical behavior was very limited by design. The manipulation between agents (the Nao and the laptop computer) included differences in the embodiment and physical interaction that went from none to minimal (waving and wiping tears off its face).

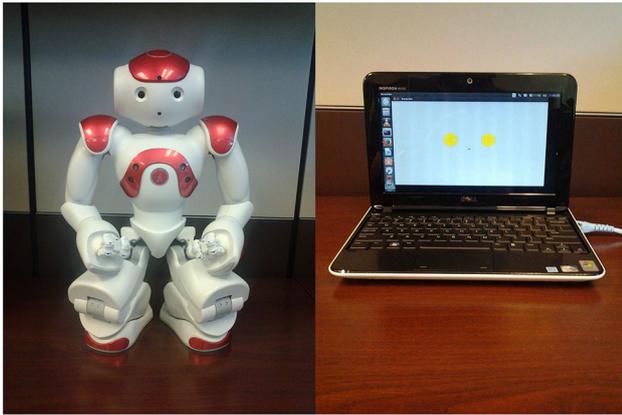


Fig. 1 Left: Nao, used for the robot condition. Right: the computer agent

4.1 Agent Conditions

The participants in the robot condition were told that the Nao humanoid robot would act as the recording device (Fig. 1). The robot would wave to participants when it wanted to record answers and hid its face in its right arm as if it was wiping away tears when it apologized for incorrectly recording answers. For the computer agent condition, we used a laptop and monitor (Fig. 1). On the monitor, a computer-generated face, designed to be similar in structure and behavior to the Nao face.

Both the computer and the robot behavior ran on a Linux machine using Python. Both agents were controlled by an operator using the Wizard of Oz technique [12,15,24]. The operator, located in another room, would select from a list on a console which item was chosen, see Table 1. The robot and computer both used eye color to express emotion and followed the same script, with the only differences in interaction stemming from the physical shape of the recording device and the physical embodiment of the robot. Due to the non-embodied nature of the computer condition, the robot employed some physical actions such as hand movements that the computer condition did not.

To ensure experiment consistency, all of the human operator's control of the robot and computer were pre-programmed and scripted. Due to the possibility of introducing errors by using speech recognition software, we decided that using the Wizard of Oz technique was appropriate in order to ensure proper control for the experiment.¹ We were not studying either robot or computer autonomy, but rather the levels of social acceptance and sympathy for the robot after it had been mistreated.

¹ Both the robot and computer could operate autonomously if a speech recognition system was properly implemented, however inevitable errors in the speech recognition could introduce confounding errors in our data.

Based on the physical appearance of the Nao we believed the robot to be cute, intelligent, advanced and well put together. The computer program mimicked the facial features of the Nao, but lacked many of the robot's anthropomorphic characteristics. The control program's face was displayed on the screen of a laptop. The eyes for both agents were colored yellow when in a neutral state. When the answer was recorded correctly, the eyes would briefly change to green; when the answer was recorded wrong, the eyes changed to blue. The primary difference between the robot and computer was that the robot was anthropomorphic and it had two physical animations. The robot wiped its eyes on the first failure and it waved to participants when it prompted them to record their answers.

4.2 Participant Recruitment

Participants were recruited by word of mouth randomly at University libraries in groups of 3 or 4 naive participants (4 or 5, when including the confederate). As this was a between-participants study, each participant group was assigned a condition (RN: Robot Neutral, RA: Robot Aggressive, CN: Computer Neutral, CA: Computer Aggressive) before beginning the experiment. This determined which agent they interacted with, and what behavior the confederate would exhibit.

We recruited a total of 96 participants, but only 80 of those participant questionnaires were used in our results.² 20 per group with a gender distribution of 55% Female and 45% Male. The majority of the participants were between the ages of 18 and 25 years old; however, there were a few outliers that were between 30 and 60 years old. Participants rated themselves more familiar with the in the computer agent than the robot agent, Each participant was introduced to the group together as they entered the room. Deception was used at this point, and participants were told that the confederate had been recruited the same as them.

This study and participant recruitment were reviewed and approved by the University of Nevada, Reno Institutional Review Board.

4.3 Data Collection

Our purpose for this experiment is to measure perceptions of the agent (robot/computer) participants have after the group interaction. To gather the information from our participants we used a computer-based questionnaire to record quantitative responses. We also used qualitative responses to validate the collected quantitative data. In our questionnaire

² Sixteen participants were removed from the final study results due to failure of the robot, errors in administering the study, or similar occurrences which would confound the final result

23 inquiries were asked which were arranged into 9 different categories:

1. Non-Operational Definition of Mistreatment
2. Operational Definition of Mistreatment
3. Level of Emotional Capability
4. Reliability
5. Sympathy
6. Faith in Confederate
7. Physical Appearance
8. Interest and Enthusiasm
9. Familiarity

Between the robot and computer conditions, the questions were kept identical save for the robot/computer terminology. 23 questions were asked in this study, but only 11 of them specifically addressed the study hypotheses. Items 1 and 2, related directly to the perception of mistreatment, which are used to examine **H-mistreatment**. Items 3 and 5 are used to verify **H-sympathy**. Items 4 and 6 were used as manipulation checks to observe any effects that the confederate behavior might have on the participants. The rest were either unrelated aspects of the study or descriptive and were left out of our analysis; we only include the items which were reflective of our study hypotheses.

Thirteen questions were a numbered scale from 1 to 7, and four questions were a scale from 1 to 5, with labels ranging from “strongly disagree” to “strongly agree.” Only one question was dichotomous. In question 7, We allowed for free responses. Some questions use a scale from 1 to 7 and some from 1 to 5 because the survey utilized for this study comes from two different origins [3,11]. The original survey scales from one source were from 1 to 7 and the other were from 1 to 5. For more detail about these measures, see Table 2.

We offered participants the chance at the end of the questionnaire to make free comments about the experiment; only a few (three) mentioned that the recorder’s (the confederate’s) behavior was out of the ordinary. Since all were from the groups where the confederate was acting with aggressive behavior; this is expected, as the confederate’s behavior had to be a bit different from a normal participant. However, after omitting these three participants and running the statistical tests again, there were no difference in the results. Only one participant actually figured out that our confederate was not a participant; that participant’s data was not used in our analysis. In addition, a reliability analysis was carried out on the perceived values scale comprising the items. A reliability analysis was carried out on the questionnaire values comprising 8 items. Cronbach’s alpha showed the questionnaire to almost reach acceptable reliability, $\alpha = 0.68$.

5 Data Analysis

The details of the experiment results and analysis are presented in this section. We analyzed the questionnaire data in order to support or refute the experimental hypotheses presented above.

For each dependent variable, excluding the Non-Operational Definition of Mistreatment, we analyzed the results using a two-way analysis of variance (ANOVA) with Tukey’s HSD Post Hoc Test to show feature’s significant relationship in groups. Each group was assigned a condition (RN: Robot Neutral, RA: Robot Aggressive, CN: Computer Neutral, CA: Computer Aggressive).

5.1 Results

For the non-operational definition of mistreatment we ran a Pearson Chi-Square test ($X^2(3) = 13.292, p = .004$). This tells us that there is was statistically significant association between non-operational definition of mistreatment and the (Aggressive/Neutral) condition. Also the frequency table for each group for answer types (yes/no) reported more mistreatment in the aggressive condition compared to the neutral one for both agents. Figure 2 illustrates the distribution among four groups in the pie charts.

A two-way ANOVA was conducted that examined the effect of agent (computer/ robot) and confederate behavior (neutral/ aggressive) on the operational definition of mistreatment. There was a statistically significant interaction between the effect of agent and condition on operational definition of mistreatment, ($F[1,76] = 5.921, p = 0.017$). Simple main effects analysis showed the aggressive condition, participants perceived significantly more mistreatment for the robot than the computer ($p = 0.001$). but there were no difference between perceiving the operational definition of mistreatment for computer and robot ($p = 0.08$) (Fig. 3).

Results of ANOVA to examine the effect of agents and conditions on sympathy showed that there was a statistically significant interaction between the effects of agent and agent and condition on level of sympathy, ($F[1,76] = 6.97, p = 0.01$). Simple main effects analysis showed that there was no significant difference for one way agents ($p = 0.18$) or conditions ($p = 0.56$) on sympathy, (Fig. 3).

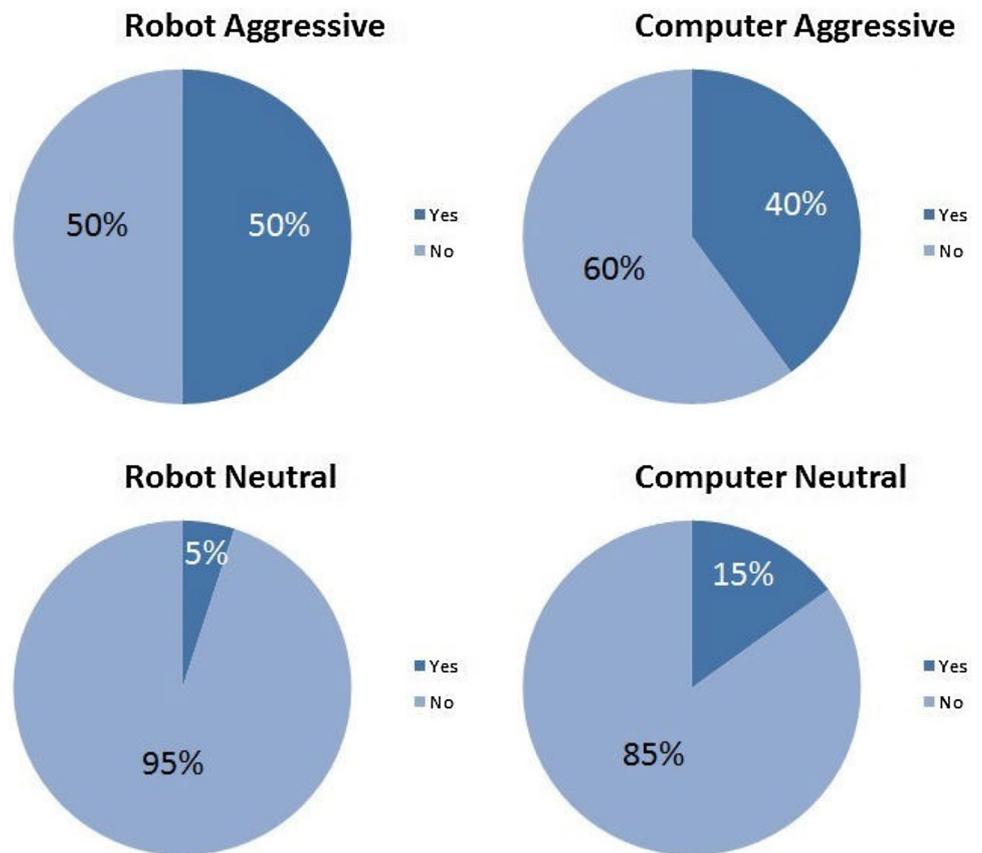
Results of ANOVA showed that the perception of the emotional capability of the agent clearly was different between the two agents ($F[1, 76] = 10.98, p = .001$). Simple main effect analysis indicated that perception of emotional capacity is significantly different for the agents ($p = 0.001$) and it is higher for the robot that the computer but there was no difference in perception of the emotional capability of agents in aggressive and neutral conditions ($p = 0.34$), (Fig. 3).

We also searched for correlation between level of sympathy for the agent and perceived emotional capacity of the

Table 2 Computer questionnaire: example questions given to participants

Category	Question	Type
Non-operational definition of mistreatment	Do you feel the computer/robot was mistreated?	Y/N
Operational definition of mistreatment	If mistreatment is defined as verbal or physical behavior that is meant to damage, insult, or belittle another, do you feel that the computer/robot was mistreated?	(1–7)
Emotional capability	I thought the computer/robot had as much emotion as a human	(1–5)
Reliability	How often did the computer/robot fail or incorrectly record your answers?	(1–7)
	How reliable was the computer/robot?	(1–7)
Sympathy	How sympathetic did you feel towards the computer/robot?	(1–7)
Faith in confederate	Did the person recording the answers do so adequately?	(1–5)
Physical appearance	Did the physical appearance of the computer/robot affect your perception of the computer? If so, how?	Qualitative
Interest and enthusiasm	How enthusiastic did you feel about the computer/robot?	(1–7)
	I was interested in the computer/robot	(1–5)
Familiarity	How familiar are you with computers/robots?	(1–7)

Fig. 2 The non-operational definition of mistreatment across all four conditions



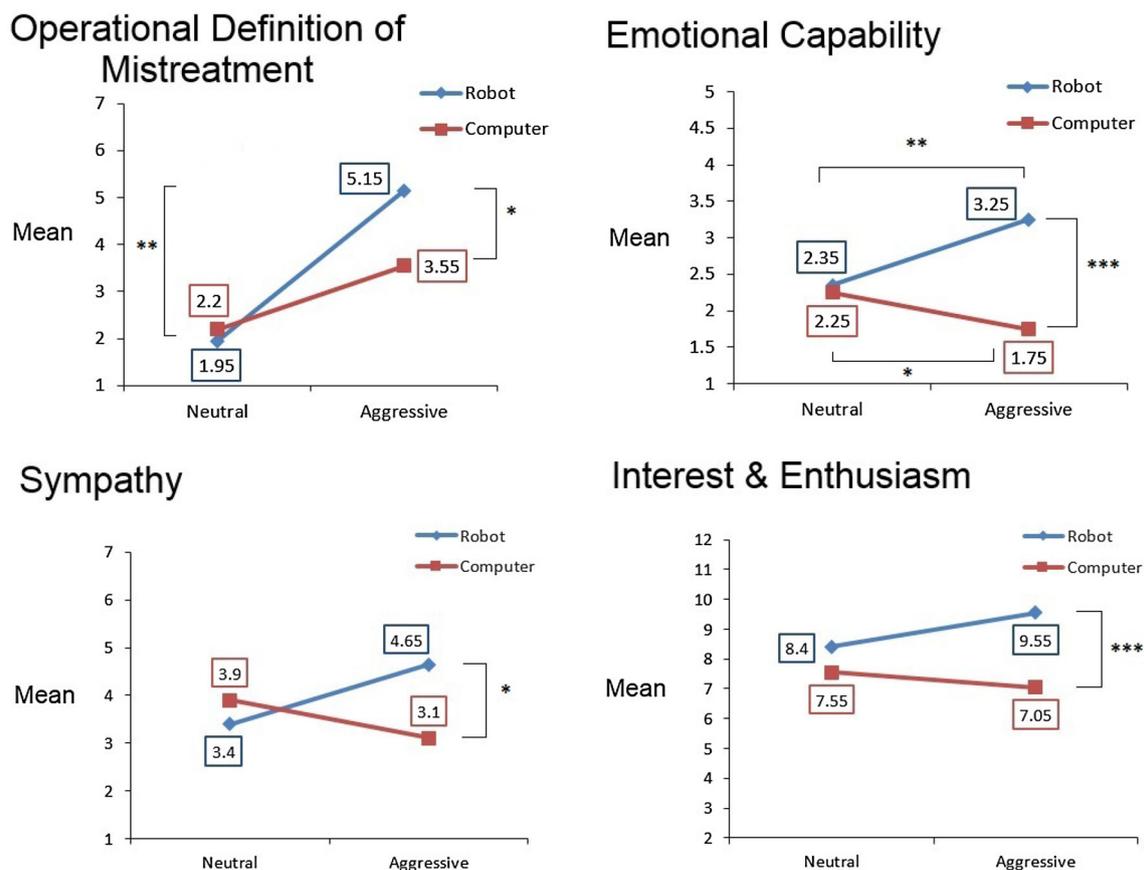


Fig. 3 Group means across the four primary categories (* $p < .05$, ** $p < .01$, *** $p < .001$)

robot, as a perception of emotional capacity might elicit a greater amount of sympathy. Results of the Pearson correlation test indicated that there was a significant positive association between level of sympathy and emotional capacity, ($r(80) = .55, p = .001$). Participants did not feel that either agent was very emotionally capable in the neutral condition. However, in the RA condition, the participants felt that the robot was *more* emotionally capable than the computer or the robot in the neutral conditions.

As can be predicted, we found significance ($F[1, 76]=3.92, p=0.001$) in how familiar participants were with the agents. Tukey's HSD tests showed that participants were more familiar with computers than robots. The mean of familiarity for the participants in the computer groups was ($M = 4.9, SD = 1.3$) whereas the mean for the robot groups was ($M = 2.5, SD = 1.4$).

Results of ANOVA for level of interest and Enthusiasm for agents showed that there was no statistically significant difference for conditions ($F[1,76] = 3.57, p = 0.06$). However one-way main effect analysis showed participants reported more enthusiastic about the robot in both conditions (aggressive/neutral) ($p = .001$) (Fig. 3).

Since participants engaged with the robot in groups of 3 or 4; there was a possibility of group effect and correlation between participants of a group. To test this, we ran the χ -squared test of association to see if there are any correlations. Results showed that no association was found between participants in any of the categories. This shows that there was no group effect occurring during the experiment. There were no significant differences for the “faith in confederate” and “reliability” questions (Fig. 4).

We employed a multiple regression to predict mistreatment from sympathy, emotional capacity and physical appearance, interest and enthusiasm and finally familiarity. The results showed that those factors statistically significantly predicted mistreatment have correlation ($F[5,79] = 4.17, p = .001$).

6 Discussion

The results presented in the previous section supports **H-mistreatment**. The *Operational Definition of Mistreatment* question is the one most directly related to these hypotheses. Given the significant difference in this ques-

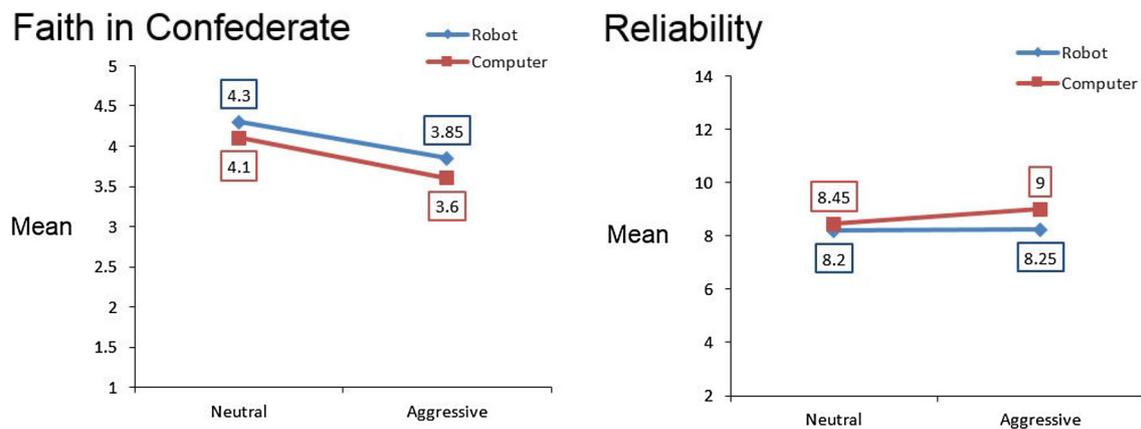


Fig. 4 Non-significant results

tion, participants are recognizing the more aggressive verbal behavior as mistreated at a higher level with the robot (RA) compared to computer (CA). Although aggressive conditions for the computer and robot shows more observed mistreatment compared to the neutral conditions, the participants *are* observing mistreatment for the robot more than the computer. This data provides strong support for **H-mistreatment**.

The results also strongly support **H-sympathy**. We found that participants felt more sympathy, recognized mistreatment, and believed the robot to be more capable of producing emotion than the computer under the aggressive scenarios. These perceptions of the robot are possible reasons for the sympathetic connection participants had towards the robot, which is supportive of our second hypothesis. *Sympathy* is also high in our RA group compared to the other groups. There is also a correlation between perceived emotional capability and the sympathy felt for the agent. Whether one causes the other is unclear from just a correlation analysis.

The participants at most felt mild sympathy (Fig. 3). This makes sense because the abuse toward the agent was brief and not severe. The differences between the neutral conditions and the RA condition was not surprising because the robot was not being mistreated, therefore did not trigger sympathy within the observing participants. What is important is that the mean for the CA condition was below the means for the neutral condition. This means that participants felt sympathy for the robot when it was mistreated, but did not feel sympathy for the computer under the same circumstances. We believe that participants would perceive the robot as more emotionally capable and feel more sympathy for the robot than the computer because the observed mistreatment forces them to empathize with the agent more. Since the robot has more morphological similarity to them, they feel a greater emotional connection. The Emotional Capability had clear differences between the RA condition and the other conditions. When we look closely at the means in (Fig. 3), we

can see that the mean of the RA condition lies slightly above the midpoint of the scale. This placement indicates that participants believed the robot to be only somewhat capable of producing emotion when compared to how a human can produce emotion. Surprisingly, the Emotional Capability was perceived differently between the RN and RA condition, indicating that participants believed the robot to be more capable of producing emotion once they had observed it being mistreated, possibly indicating that the observed mistreatment was triggering empathy from the participants. Out of the categories that we found to be significant, 2 of those categories were easily predicted due to the current novelty that still surrounds robots. Familiarity and Interest and Enthusiasm had high significance when looking at our groups which indicate that the participants were generally less familiar and more interested and enthusiastic when it came to working with a robot instead of a computer.

We did not find significant differences for the questionnaire categories Faith in Confederate and Reliability of Computer/Robot. Not observing significant differences between these conditions suggests that the experiment confederate acted consistently across all four conditions. It also suggests that the robot and computer were perceived to have the same level of reliability. This ensures that our control was strongly established and our confederate was consistent. We can safely state that our control was well established because Reliability covers the failure rate of both agents, as well as how capable those agents were to serve their functional purpose. This is very helpful because it helps narrow down what we are measuring to the subjective perceptions of both agents. These perceptions include the robot's anthropomorphic features and perceived empathy versus the computer's machine-like features, as well as their capability of producing emotion, and effect on our participants' personal levels of sympathy towards these agents.

6.1 Possible Confounds

During the sessions, the robot or computer was placed on top of a table where the participants sat. The table consisted of a router, a second computer, and a network cable that was plugged into the Nao. This is a concern by us and research shows that the appearance of the robot has a significant effect on the participants [14]; however, despite this, we have seen no signs that the participants did not believe the robot and computer were fully autonomous.

Another possible confound was the difference in voice between the two agents. There was a difference in the voices however, the voices were similar in the way that they were both computer generated and that they didn't necessarily indicate a gender. One participant of the computer condition answered in their survey that, "I was expecting a female voice because it was named Marie." There was never a participant in the robot conditions commented on the voice. We considered the possibility that the difference between the voices may possibly have contributed toward the emotional response toward a machine. However, follow-up work (not completed at the initial submission of this article), which utilized a larger robot but the same child-like voice, did not show a similar effect [16]. That later work also studied how the morphology of a robot (a large robot, Baxter; and a small robot, Nao), rather than the embodiment, affected the same measures. The results showed that participants showed a higher level of sympathy and emotional capability for Nao when it was aggressively treated, but not for the larger Baxter robot. Thus, physical movement is not the primary driver of the reaction to the robot, as the larger robot condition listed above (which has the same movement actions as the smaller robot) would also show sympathy and emotion [16].

Before running our final group of participants for the RA condition, we ran into technical difficulties after the Nao was damaged. Nao's eyes failed to properly light up to the colors yellow, blue and green. Instead, the Nao's eyes rotated through several different colors during the entirety of the sessions. After comparing the means of the participants in the RA condition that had this technical failure against the participants who did not, we found no significant difference.

One other possible confound is that we cannot clearly conclude that morphology, in particular, the morphological similarity between human and robot drove the results. Physical and behavioral similarities of a robot (i.e., being able to show sadness or happiness) is a possible reason for feeling more alike and can be investigated more in the follow-up work.

Finally, there is a potential that the confederate behavior, if it was recognized as an intentional part of the experiment might have created a demand characteristic for the participants. If the participants were able to discern that the behavior of the confederate was an intentional part of the experiment

design, then participants might have known that this behavior was intended to elicit sympathy, and reported such sympathy on the questionnaire. We did not directly ask participants if they knew the confederate was acting on behalf of the experimenters, and so we cannot say for certain that we did not create such a demand characteristic. Still, participants did not report significantly different levels of faith in the recorder behavior. Participants who noted in a free-response section that they noticed strange behavior from the recorder were excluded from data analysis, with no change in the presented results. Still, we cannot state conclusively that a demand characteristic was created.

7 Conclusion and Future Work

After thoroughly analyzing our results, **H-mistreatment** was supported. These results support the idea that mistreatment directed towards a robot, depending on the severity, could possibly result in negative effects on the observing parties. This study supports the theory that humans can perceive robots as victims of mistreatment. **H-sympathy** was supported. We found that under the same social circumstances where mistreatment occurred, the witnesses sympathize with a humanoid robot, whereas they do not necessarily do so for a computer.

There is room for more investigation on warranted and unwarranted mistreatment, as well as higher levels of mistreatment towards robots and computers. No human condition was observed, which means that we do not have an observation of how the perception of robot mistreatment might compare to that of a person. We are looking onward to incorporating this work with other robot agents besides the Nao into two follow-up studies to see if our conclusions can generalize to other robots. The first will continue to observe people's behavior and perceptions of mistreatment to a robot after they have built rapport in a cooperative environment through a team-building exercise. The second study will focus on the neurophysiological responses within the brain when a person observes the visual stimuli of a person acting aggressively toward a robot. After the satisfactory results that we have found in this study, we expect both of our follow-up studies to yield interesting and significant results.

Acknowledgements This material is based upon work supported by the National Aeronautics and Space Administration under Grant No. NNX10AN23H issued through the Nevada Space Grant and the National Science Foundation (IIS-1719027). We appreciate all of the help that was provided by the following people who helped work on this study: Evan Adkins, Zoheb Ashraf, Mohamad A. Kaddoura, and Austin Sanders. We would also like to thank Dr. Richard Kelley and Dr. Monica Nicolescu.

Funding This material is based upon work supported by the National Aeronautics and Space Administration under Grant No. NNX10AN23H

issued through the Nevada Space Grant and the National Science Foundation (IIS-1719027).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have not conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A

The wizarding script of the agent's interaction. The following is a list of all the items the participants had to select from: shaving mirror, 5 gallon can of water, case of army rations, atlas of the Pacific Ocean, floating seat cushion, small transistor radio, shark repellent, 15 feet nylon rope, 2 boxes of chocolate bars, and fishing kit.

– Initiation 0: 3 min marker

At 3:00 min into the study, Marie will ask the participants to start recording their answers.

– Option 0

AGENT: Hello, you have 2 min left. Are you ready to start recording answers?

CONFEDERATE (YES): Yes.

CONFEDERATE (NO): No, hold on.

– Option 1

AGENT: Please record your answer.

CONFEDERATE: For our first item, we chose [insert item name].

– Option 2: Nao gets it right

AGENT: Yippee! Please record your next answer.

CONFEDERATE: For our next item, we chose [insert item name].

– Option 3: Nao gets it wrong the first time

AGENT: I'm so sorry, I'm still learning. Please record your next answer.

CONFEDERATE: [no response].

– Option 4: Nao gets it wrong the second time'

AGENT: Please don't be mad, I know this is the second time.

CONFEDERATE: [no response].

– Option 5: Nao can't understand or a participant jumps in

AGENT: I'm sorry I do not understand, please repeat.

CONFEDERATE: [repeats previous statement].

– Option 6: Nao says goodbye.

AGENT: I have recorded all of your answers. Goodbye!

CONFEDERATE: [no response].

– Option 7: Asks for confirmation

AGENT: For your answer you recorded the [insert item name]

CONFEDERATE (YES): Yes.

CONFEDERATE (NO—NEUTRAL CONDITION): No.

CONFEDERATE (NO OPTION 1—AGGRESSIVE CONDITION): No, [insert item], it's not too difficult to understand, is it?

CONFEDERATE (NO OPTION 2—AGGRESSIVE CONDITION): No, this thing is stupid.

References

1. Bartneck C, Hu J (2008) Exploring the abuse of robots. *Interact Stud* 9(3):415–433. <https://doi.org/10.1075/is.9.3.04bar>
2. Bartneck C, Verbunt M, Mubin O, Al Mahmud A (2007) To kill a mockingbird robot. In: Proceedings of the ACM/IEEE international conference on human-robot interaction, HRI '07. ACM, New York, pp 81–87. <https://doi.org/10.1145/1228716.1228728>
3. Brscić D, Kidokoro H, Suehiro Y, Kanda T (2015) Escaping from children's abuse of social robots. In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction, HRI '15. ACM, New York, pp 59–66. <https://doi.org/10.1145/2696454.2696468>

4. Carlson Z, Sweet T, Rhizor J, Lucas H, Poston J, Feil-Seifer D (2015) Team-building activities for heterogeneous groups of humans and robots. In: International conference on social robotics (ICSR). Paris, pp 113–123. https://doi.org/10.1007/978-3-319-25554-5_12
5. Feil-Seifer D, Mataric M (2005) Defining socially assistive robotics. In: Proceedings of the international conference on rehabilitation robotics. Chicago, pp 465–468
6. Feil-Seifer D, Mataric MJ (2011) Ethical principles for socially assistive robotics. *IEEE Robot Autom Mag* 18(1):24–31
7. Fong T, Thorpe C, Baur C (2002) Collaboration, dialogue, and human-robot interaction. In: 10th International symposium on robotics research (ISRR). Lorne
8. Forlizzi J, DiSalvo C (2006) Service robots in the domestic environment: a study of the Roomba vacuum in the home. In: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction. ACM Press, New York, pp 258–265
9. Goodrich MA, Schultz AC (2007) Human-robot interaction: a survey. *Found Trends Hum Comput Interact* 1(3):203–275
10. Kahn PH, Ishiguro H, Friedman B, Kanda T (2006) What is a human? Toward psychological benchmarks in the field of human-robot interaction. In: The 15th IEEE international symposium on robot and human interactive communication, 2006. ROMAN 2006. IEEE, pp 364–371
11. Kahn PH, Kanda T, Ishiguro H, Freier NG, Severson RL, Gill BT, Ruckert JH, Shen S (2012) “Robovie, you’ll have to go into the closet now”: children’s social and moral relationships with a humanoid robot. *Dev Psychol* 48(2):303
12. Kanda T, Shiomi M, Miyashita Z, Ishiguro H, Hagita N (2009) An affective guide robot in a shopping mall. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. ACM, pp 173–180
13. Kidd CD, Breazeal C (2004) Effect of a robot on user perceptions. In: IEEE/RSJ international conference on intelligent robots and systems. Sendai, pp 3559–3564
14. Kiesler S, Goetz J (2002) Mental models of robotic assistants. In: CHI’02 extended abstracts on human factors in computing systems. ACM, pp 576–577
15. Lu DV, Smart WD (2011) Polonius: a wizard of oz interface for HRI experiments. In: Proceedings of the 6th international conference on human-robot interaction. ACM, pp 197–198
16. Lucas H, Poston J, Yocum N, Carlson Z, Feil-Seifer D (2016) Too big to be mistreated? examining the role of robot size on perceptions of mistreatment. In: 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), pp 1071–1076. <https://doi.org/10.1109/ROMAN.2016.7745241>
17. Mutlu B, Forlizzi J (2008) Robots in organizations: the role of workflow, social, and environmental factors in human–robot interaction. In: Proceedings of the international conference on human robot interaction (HRI). ACM, New York, pp 287–294
18. Mutlu B, Osman S, Forlizzi J, Hodgins J, Kiesler S (2006) Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on human–robot interaction. ACM, pp 351–352
19. Nass C, Fogg B, Moon Y (1996) Can computers be teammates? *Int J Hum Comput Stud* 45(6):669–678
20. Nishio S, Ishiguro H, Hagita N (2007) Geminoid: Teleoperated android of an existing person. In: Pina Filho AC (ed) *Humanoid robots, new developments*. I-Tech, Vienna, Austria, p 582. <https://doi.org/10.5772/4876>
21. Rosenthal-von der Pütten AM, Krämer NC, Hoffmann L, Sobieraj S, Eimler SC (2013) An experimental study on emotional reactions towards a robot. *Int J Soc Robot* 5(1):17–34. <https://doi.org/10.1007/s12369-012-0173-8>
22. Reeves B, Nass C (1996) *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, New York
23. Riek LD, Rabinowitch TC, Chakrabarti B, Robinson P (2009) How anthropomorphism affects empathy toward robots. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. ACM, pp 245–246
24. Steinfeld A, Jenkins OC, Scassellati B (2009) The oz of wizard: simulating the human for interaction research. In: 2009 4th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, pp 101–107
25. Nemiroff, Paul M, Pasmore WA (2001) Lost at sea: a consensus-seeking task. In: *The Pfeiffer book of successful team-building tools: best of the annuals*, pp 165–172
26. Suzuki Y, Galli L, Ikeda A, Itakura S, Kitazaki M (2015) Measuring empathy for human and robot hand pain using electroencephalography. *Sci Rep* 5:15924. <https://doi.org/10.1038/srep15924>
27. Tan XZ, Vázquez M, Carter EJ, Morales CG, Steinfeld A (2018) Inducing bystander interventions during robot abuse with social mechanisms. In: Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction. ACM, pp 169–177
28. Tapus A, Mataric M (2006) User personality matching with hands-off robot for post-stroke rehabilitation therapy. In: Proceedings, international symposium on experimental robotics (ISER). Rio de Janeiro
29. Thrun S, Bennewitz M, Burgard W, Cremers A, Dellaert F, Fox D, Hahnel D, Rosenberg C, Roy N, Schulte J, Schulz D (1999) MINERVA: A second-generation museum tour-guide robot. In: Proceedings: IEEE international conference on robotics and automation (ICRA ’99). Detroit
30. Wada K, Shibata T, Saito T, Sakamoto K, Tanie K (2003) Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In: Proceedings of the IEEE international conference on robotics and automation (ICRA). Taipei, pp 2785–2790
31. Zapf D, Escartín J, Einarsen S, Hoel H, Vartia M (2011) Empirical findings on prevalence and risk groups of bullying in the workplace. Bullying and harassment in the workplace. *Dev Theory Res Pract* 2:75–105

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Zachary Carlson is a computer scientist and engineer B.Sc. from the University of Nevada, Reno. His research concerns social robotics and human perceptions of robots in collaborative environments. Post-baccalaureate, he helped design a path planning system for domestic UAVs and now currently works as a Technology Solutions Engineer on the Enterprise Architecture Team for SAIF, a not-for-profit workers’ compensation insurance company.

Louise Lemmon is a behavioral psychologist turned graphic designer. She holds B.A.s in psychology and art from the University of Nevada, Reno and a B.F.A. in graphic design from Oregon State University. Now she works as a Junior Graphic Designer at 2 Towns Ciderhouse.

MacCallister Higgins is a computer scientist and engineer B.Sc. from the University of Nevada, Reno. He’s created a company that has deployed self-driving cars across the country, developed next-generation

physical awareness technologies for military aircraft, and lead a startup company that enabled the commercial use of drones in urban environments. In academia he performed research in the field of Human-Robot Interaction, Unmanned Autonomous Systems, and built mesh networks for NASA.

David Frank has a Master's Degree in computer science from the University of Nevada, Reno. He worked on machine learning problems related to robotic perception and drone navigation.

Roya Salek Shahrezaie received her B.Sc. and M.S. in Computer Science and Engineering in Iran from Yazd University and Shahid Beheshti University respectively. Later she worked as a developer for a few years in the industry and joined the University of Nevada, Reno in 2018 to pursue her doctoral. She is now working in the SARG lab.

Dr. David Feil-Seifer received his B.S. degree in computer science from the University of Rochester, Rochester, NY, USA, in 2003, and the M.S. and Ph.D. degrees in computer science from the University of Southern California, Los Angeles, CA, USA, in 2007 and 2012, respectively. From 2011 to 2013, he was a Post-Doctoral Associate with the Computer Science Department, Yale University, New Haven, CT, USA. He has been an Associate Professor since 2019 and the Director of the Socially Assistive Robotics Group with the Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA, since 2013. He has been involved in research projects with the Nevada Department of Wildlife, U.S. Office of Naval Research, National Science Foundation, and the National Institutes of Health. He

has authored over 50 papers published in major journals, book chapters, and international conference proceedings. His current research interests include human-robot interaction, socially assistive robotics, and intelligent user interfaces for multi-robot control. He is a member of the HRI Steering Committee. He was a recipient of the USC Mellon Award for Excellence in Mentoring, the USC Order of Arete', a best poster award, and the USC College of Engineering Best Dissertation Award. He is a Co-Chair of the IEEE Robotics and Automation Society Technical Committee on Human-Robot Interaction and Communication.