

## This robot stinks! Differences between perceived mistreatment

Zachary Carlson

zack@nevada.unr.edu

Robotics Research Lab, University of Nevada, Reno

Louise Lemmon

louise@nevada.unr.edu

Robotics Research Lab, University of Nevada, Reno

MacCallister Higgins

mac@nevada.unr.edu

Robotics Research Lab, University of Nevada, Reno

David Frank

davidfrank@nevada.unr.edu

Robotics Research Lab, University of Nevada, Reno  
and

Dr. David Feil-Seifer

dave@cse.unr.edu

Robotics Research Lab, University of Nevada, Reno

### **Intended Type of Contribution:**

Research Article

### **Keywords:**

human-robot interaction

human-robot cooperation

perception

mistreatment

quantitative field study

service robots

ethics

### **Relationship Between Most Closely Related Prior Papers:**

None of the authors have any related papers to this work. This paper was originally submitted to HRI 2015 but it was not accepted and we have revised the paper according the feedback given by the reviewers.

### **Role of Authors:**

Author 1, 2, and 5 are responsible for writing and editing the paper. All authors were involved in the experiment design. Author 3 was the confederate and Author 4 was the programmer behind the computer.

### **(Optional) Suggested Reviewers:**

Peter H. Kahn, Jodi Forlizzi, Tony Belpaeme, Julie Shah, Christoph Bartneck, Leila Takayama, Cindy Bethel

### **Authors Willing to Present:**

Zachary Carlson, Louise Lemmon, David Feil-Seifer

# This robot stinks! Differences between perceived mistreatment of robot and computer partners

Zachary Carlson

Robotics Research Lab, University of Nevada, Reno

Louise Lemmon

Robotics Research Lab, University of Nevada, Reno

MacCallister Higgins

Robotics Research Lab, University of Nevada, Reno

David Frank

Robotics Research Lab, University of Nevada, Reno

and

David Feil-Seifer

Robotics Research Lab, University of Nevada, Reno

---

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 animacy of such agents. Mistreatment of robots (or computers) by co-workers might provoke such telling reactions. The purpose of this study was to discover if people perceived mistreatment directed towards a robot any differently than toward a computer. This will provide some understanding of how people perceive robots in collaborative social settings.

We conducted a between-subjects study with 80 participants. Participants worked cooperatively with either a robot or a computer which acted as the “recorder” for the group. A confederate either acted aggressively or neutrally towards the “recorder.” We hypothesized that people would not socially accept mistreatment towards an agent that they felt was intelligent and similar to themselves; that participants would perceive the robot as more similar in appearance and emotional capability to themselves than a computer; and would observe more mistreatment. The final results supported our hypothesis; the participants observed mistreatment in the robot, but not the computer. Participants felt significantly more sympathetic towards the robot and also believed that it was much more emotionally capable.

Keywords: Human-robot interaction, mistreatment, perception, human-robot cooperation

---

---

Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work’s authorship and initial publication in this journal.

## 1. Introduction

Robots are quickly becoming part of people's everyday lives yet we lack an understanding of what social roles robots might play (Feil-seifer & Matarić, 2011). Robots that resemble humans and display social intelligence are being deployed in work, home, and care settings (Feil-Seifer & Matarić, 2005). There is a large and growing volume of HRI studies showing positive robot behavior and positive human interaction with robots (Wada, Shibata, Saito, Sakamoto, & Tanie, 2003; Tapus & Matarić, 2006; Fong, Thorpe, & Baur, 2002; Kanda, Shiomi, Miyashita, Ishiguro, & Hagita, 2009). 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 we don't know what kind of impact this can have. Thus, it should be a priority to study the full relationship between humans and robots - not just positive interactions.

One example of such behavior is workplace bullying. Workplace bullying is not an uncommon sight between people in office settings. It can be something as subtle as a backhanded complement, to humiliating a colleague or more aggressive behavior (Zapf, Escartín, Einarsen, Hoel, & Vartia, 2011). Such behavior could result in negative psychological effects not just for those that are bullied, but also for those observing the bullying (Vartia, 2001).

It can be just as common for machines to receive this same kind of treatment. A copy machine, for example, might be physical 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? How about a robot that was capable of exhibiting an emotional response?

Given that Human-Robot Interaction (HRI) and Human-Computer Interaction (HCI) can emulate Human-Human Interaction (HHI) (Reeves & Nass, 1996), it is conceivable that a similar reaction to observed mistreatment might occur between humans and robots. One can imagine 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 OK to kick a jammed copy machine, but is it also OK to kick a robotic dog that ran into your leg? What about a small humanoid robot that resembles a child? The different levels of embodiment may have significantly different effects on the people's interactions with the agents and their perceptions of them. By quantifying that social dividing line for the acceptance of targeted mistreatment towards robots, this study will potentially contribute a piece of research towards discovering that answer.

In this paper, we compare participants' reactions to verbally abusive behavior (not physically abusive behavior) toward a computer and a robot. First, we will explore prior research of how people in alone or in groups treat (or may mistreat) robots. Then, we present a controlled study to examine how verbal mistreatment might affect collaborators with a robot. We then present results supporting our hypotheses that people will perceive mistreatment of a robot more than they will for a computer, given the same treatment due to their differences in visual form. These results contribute to our understanding of how people perceive robots in a cooperative environment.

## 2. Background

Technology is more and more becoming an integral part of our lives. The Roomba, a robot that can autonomously vacuum rooms in a house or office environments, is an example of this exact type of integration (Goodrich & Schultz, 2007). Studies have demonstrated that a family will adjust its behavior to accommodate the operation of such a robot (Forlizzi & DiSalvo, 2006). 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 perceptions of robots is an interesting topic of study, thoroughly explored using a variety of robot scenarios, through observing human interaction with those robots (Thrun et al., 1999; Kanda et al., 2009; Riek, Rabinowitch, Chakrabarti, & Robinson, 2009; Mutlu, Osman, Forlizzi, Hodgins, & Kiesler, 2006).

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 (Mutlu & Forlizzi, 2008).

This difference in treatment of the robot 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 (Zapf et al., 2011), how would mistreatment of a robot by a human co-worker affect other people in that environment?

Given that robots are becoming consistently more similar in appearance to human beings (Nishio, Ishiguro, & Hagita, 2007), there is an understandable confusion as to the role of the robot. Sharkey and Sharkey argue that robots may suborn bad parenting through entrained indifference (Sharkey & Sharkey, 2010). Kahn, et al., developed a set of benchmarks to evaluate how human-like the robot's actual behavior was (Kahn, Ishiguro, Friedman, & Kanda, 2006). 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 children will be exposed to robots. 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 (Kahn et al., 2012). 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.

Reeves and Nass 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 (Reeves & Nass, 1996). This effect means that directly asking people about the moral standing of robots (as done in the Kahn studies above) might miss these implicit changes in attitude and behavior. Nass 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 (Nass, Fogg, & Moon, 1996). While prior work has examined the effects of perceived animacy of robots, we dispute the notion that the "ultimate test for the life-likeness of a robot is to kill it" (Bartneck, Verbunt, Mubin, & Al Mahmud, 2007). We propose employing a human-robot collaboration scenario. 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 acceptance of directed mistreatment towards robots is important for the continued integration of robots into our daily lives. In the next section, we present a study that aims to contribute toward this goal.

### 3. Experiment Design

This section will present an experiment that examines social interaction with robots. In this experiment, participants observe the mistreatment of either a robot or a computer by an experiment confederate. Participant reaction was measured through surveys to determine if there is a difference in observer opinion regarding 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 (*Lost at Sea exercise from Training Manager Success Strategies*, 2014). The items ranged from food supplies to survival tools. The participants were told that they only had enough space in the rubber life craft for 5 out of 10 items and to discuss as a group which ones to take.

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-minute time limit for discussing which items to take. At the three minute 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 setup a scenario where a confederate could be observed interacting with the agent.

One of these participants was, in fact, 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 (see Table 1).

Table 1: Robot and Computer Scripted Responses for all Conditions. Please see Appendix A for a more detailed explanation of the script used by the wizard.

Condition	Responses
Recorded the answer correctly	Yipee! Please record your next answer.
Recorded the answer incorrectly the first time	I’m sorry, I’m still learning.
Recorded the answer incorrectly the second time	I’m so sorry, I know this is the second time!

At this point, the main experiment manipulation occurred. For half of the groups, the confederate would react neutrally toward the agent (control group). For the other half, the confederate would act aggressively toward the agent (experiment group). 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 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 inches tall. His behavior throughout each group was scripted a priori (see Appendix A); which included actions such as: speaking slowly as if he was irri-

tated 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. Instead, he simply sat upright at the table with the answers the group provided on a piece of paper in his hands at the top of the table. 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.

In order to maintain consistency and the inner validity of our study, the confederate had an established set of insults to use for aggressive behavior. 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.

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

We used a 2x2 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 participant's reactions and perceptions of the agent. Our hypotheses were as follows:

**H1:** After observing a humanoid robot being mistreated in a social setting, people will accept the mistreatment of the robot due to: potential apathy in regards to the robot's physical appearance, functional capability, or ability to feel emotion.

**H2:** There will be a significant difference in the accepted levels of mistreatment between the robot and a computer which may be a result of a possible sympathetic connection towards the robot for varying reasons.

The first hypothesis describes a scenario where the participants are accepting the mistreatment directed towards the robot due to their perceptions that it lacks functional capability, inability to feel emotion, and the robot's physical appearance. The second hypothesis describes an opposing scenario, where participants are not accepting of the mistreatment directed towards the robot because the embodiment of the robot versus the computer results in a possible sympathetic connection.

The Nao robot was selected for its anthropomorphic features, 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. The face of the Nao robot was mimicked on the computer's screen to control for the facial features used to evoke engagement and emotional responses from the participants when it was interacting with the group (Kidd & Breazeal, 2004). It was small enough to be placed on top of the table like the laptop computer. Since this is a humanoid robot versus a basic laptop computer, the results should act as a good predictor of what we can expect would happen as the robot agents become more human-like.

The Nao robot generally seemed to be an incapable robot with a small range of interactions available for use. This was not an issue and appears to be a good match for the computer because the computer is completely incapable of physical interaction. The manipulation between agents (the

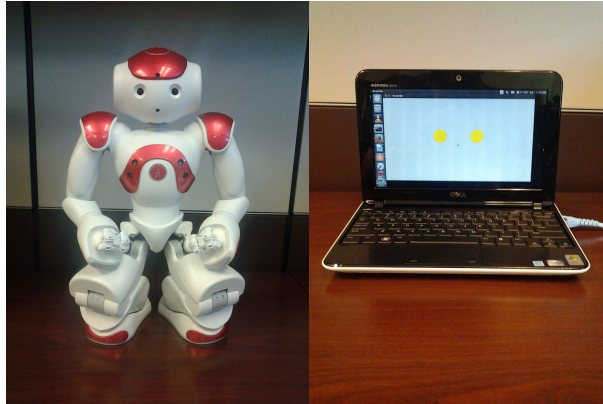


Figure 1. Left: NAO, used for the robot condition. Right: the computer agent.

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).

The flow of the conversation between the agent and the confederate stayed the same throughout the study, except for the main manipulation of how the confederate reacted when the agent failed. The confederate behavior had two levels (neutral and aggressive). The confederate in the first level would treat the agent neutrally when it had failed to record the correct answer. In the second level the confederate took on a more aggressive attitude towards the agent after it would fail to record the answer, belittling the device and expressing the lack of usefulness.

## 4. Methods

In this section, we will describe the implementation details of the experiment described above. This will include the robot and computer control behavior, participant recruitment, and data collection.

### 4.1 Agent Conditions

The participants in the robot conditions observed the NAO humanoid robot as the recording device (see Figure 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 (see Figure 1). On the monitor, a computer-generated face, designed to be similar in structure 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 (Steinfeld, Jenkins, & Scassellati, 2009; Lu & Smart, 2011; Kanda et al., 2009). 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.

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<sup>1</sup>. We

<sup>1</sup>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.

were not studying either robot or computer autonomy, but rather the levels of social acceptances and sympathy for the robot after it had been mistreated.

#### 4.2 Participant Recruitment

Participants were recruited by word of mouth at University libraries in groups of 3 or 4 (4 or 5 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 surveys were used in our results, 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 were significantly more familiar with the computer than the robot, where the mean of familiarity for the participants in the computer groups was 4.95 (out of 7) where as the mean for the robot groups were 2.53 (see Figure 2). 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 was reviewed and approved by the University of Nevada, Reno Institutional Review Board.

#### 4.3 Data Collection

Table 2: Computer Survey

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)

We used a computer survey to record quantitative responses. We used qualitative responses to validate collected quantitative data. We asked 23 questions which were scored into 9 different categories. Between the robot and computer conditions, the questions were kept identical save for the robot/computer terminology. Thirteen questions were on a scale of 1 to 7 and four questions



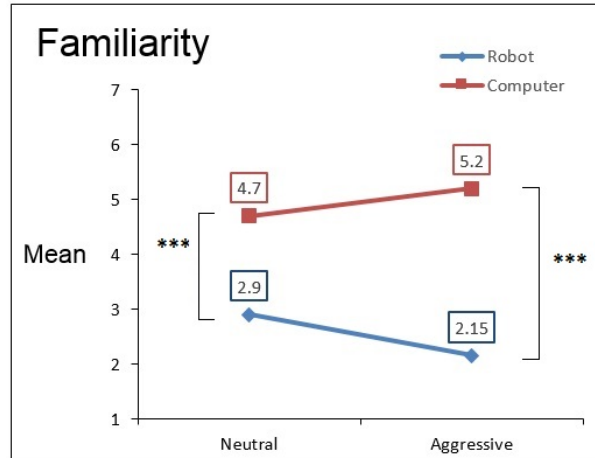


Figure 2. Question: How familiar are you with computers/robots? (\*\*\*: $p < .001$ )

were on a scale of 1 to 5. Only one question was a yes or no. We did use a qualitative question for the seventh category.

In an effort to understand how mistreatment of a robot can affect the people observing it, we measured 80 participant responses in 9 different categories: Non-Operational Definition of Mistreatment, Operational Definition of Mistreatment, Level of Emotional Capability, Reliability, Sympathy, Faith in Confederate, Physical Appearance, Interest and Enthusiasm, and Familiarity. For more detail about these measures, see Table 2.

## 5. Results

The details of the experiment results and analysis are presented in this section. We analyzed the survey data in order to support or refute the experimental hypotheses presented above. For each independent variable, we analyzed the results using an analysis of variance (ANOVA) and followed up with Tukey’s HSD test to establish significant pair-wise relationships. The groups were assigned a condition (RN: Robot Neutral, RA: Robot Aggressive, CN: Computer Neutral, CA: Computer Aggressive). Some participants had suspicions during their survey but only one participant actually figured out that our confederate was not a participant and that participant’s data was thrown out.

We observed a significant effect with the operational definition of mistreatment ( $F[3, 76] = 14.963, p < .001$ ). Pairwise tests showed that participants in the RA condition observed more mistreatment than the CA condition ( $p < .05$ ), see Figure 3.

The Non-Operational Definition of Mistreatment resulted in significance ( $F[3, 76] = 5.048, p < .01$ ). The pairwise tests showed that the RA condition vs the RN ( $p < .01$ ) and CN ( $p < .05$ ) conditions were also significant, see Figure 4

Both the non-operational and operational definitions of mistreatment resulted in high significance. These differences show that when the robot was being mistreated, participants were able to recognize this mistreatment before and after they were given the definition of mistreatment. These results indicate that participants were not able to identify that the computer was being mistreated (accepting the mistreatment) and they were able to identify that the robot was being mistreated, hinting at the possibility that they did not accept that form of mistreatment towards the robot in a social setting.

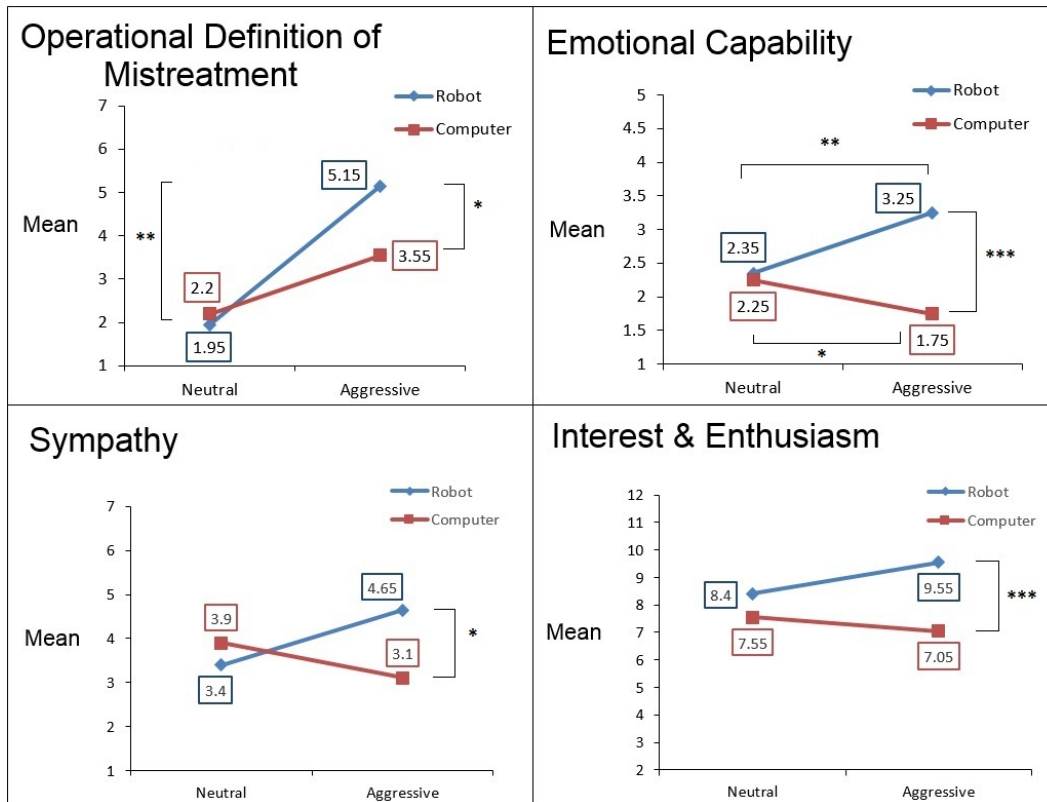


Figure 3. Group means across the four primary categories. (\*:  $p < .05$  \*\*:  $p < .01$  \*\*\*:  $p < .001$ )

The ANOVA test for the level of sympathy participants had towards the agents resulted in a significance ( $F[3, 76] = 3.047, p < .05$ ). Pairwise comparisons showed that participants felt more sympathy for the robot in the aggressive condition than the computer in the aggressive condition ( $p < .05$ ), see Figure 3.

The perception of the emotional capability of the agent resulted in clear significance ( $F[3, 76] = 8.743, p < .001$ ). The participants felt that the robot in the aggressive condition had more emotional capability than the neutral conditions ( $p < .01$ ) or the computer in the aggressive condition ( $p < .001$ ), see Figure 3.

The participants' sympathy for the agent and their belief in the emotional capability of the agent seem to be related. Again, the participants did not feel that either agent was very emotionally capable. However, in the robot aggressive conditions, the participants felt that the robot was *more* emotionally capable than the computer or the robot in the neutral conditions.

The participants reported more Interest and Enthusiasm towards the agents in all conditions ( $F[3, 76] = 6.327, p < .001$ ). The pairwise tests showed that participants observing the robot aggressive showed more interest and enthusiasm than in the computer aggressive condition ( $p < .001$ ), see Figure 3.

As can be predicted, we found a significance ( $F[3, 76] = 20.991, p < .001$ ) in how familiar participants were with the agents. The Tukey tests showed that participants were more familiar with computers than robots, see Figure 2.

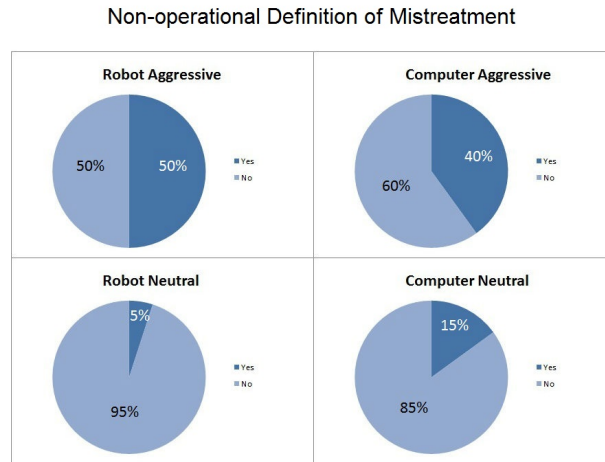


Figure 4. The Non-operational Definition of Mistreatment across all four conditions.

Out of the 6 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 indicates that the participants were generally less familiar and more interested and enthusiastic when it came to working with a robot instead of a computer.

There were no significant differences for the Faith in Confederate, Reliability, and Physical Appearance questions (Figure 5)

When participants noted that the physical appearance of the agent influenced their perception, we found that they believed the robot to be cute, intelligent, advanced and well put together. The following examples are all from the RA condition:

- “Yes, the physical appearance of the robot affected my perception of it. It was so small and cute that I couldn’t help but laugh when it would get excited.” -Participant A
- “It was pretty cute. Maybe it effected[sic] it a little, I probably wouldn’t have felt any sympathy for the robot if it just looked like a hunk of metal.” - Participant B
- “Because it was very well put together I had more faith in the robot’s competency.” - Participant C
- “Yes, it was really cool looking! Figured it used pretty advanced technology” -Participant D
- “Yes, it seemed more human like rather than a computer. It was more of a figure that I could associate with something living.” - Participant G

## 6. Discussion

The results presented in the previous section fail to support hypothesis #1, “After observing a humanoid robot being mistreated in a social setting, people will accept the mistreatment of the robot due to: potential apathy in regards to the robot’s physical appearance, functional capability, or ability to feel emotion.” There is no data suggesting that participants were apathetic and accepting of the mistreatment directed towards the robot.

The results largely support hypothesis #2, “There will be a significant difference in the accepted levels of mistreatment between the robot and a computer which may be a result of a possible sympathetic connection towards the robot for varying reasons.” In this section, we will discuss these

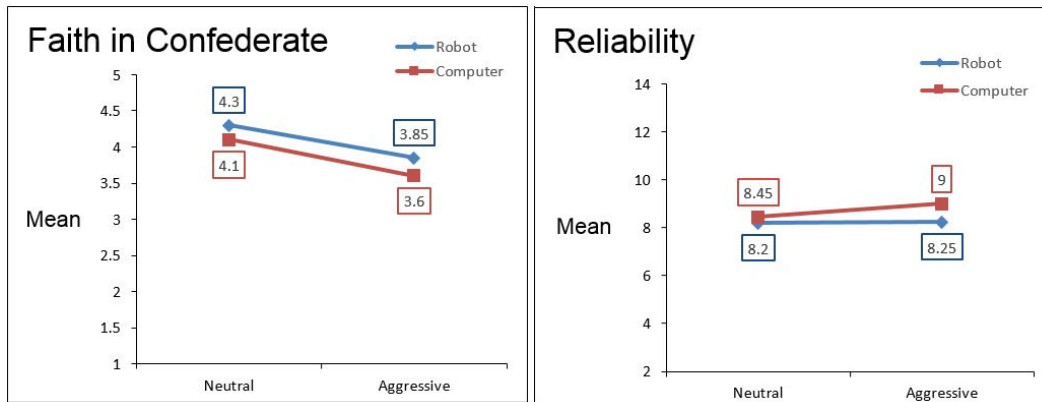


Figure 5. Non-significant results.

results in more detail.

The goal of this paper and experiment was to investigate the possible perceptions of mistreatment towards a humanoid robot in a social setting. Our second hypothesis was strongly supported by this data, while the opposing hypothesis was not. We found that participants felt 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 the possible reasons for the sympathetic connection participants had towards the robot which is supportive of our second hypothesis.

The Operational Definition of Mistreatment question is the one most directly related to these hypotheses. The significance of this result given a standardized definition of mistreatment, means that participants observed more mistreatment in the robot condition than in the computer condition. It is telling, however, that both aggressive conditions (computer and robot), that the participants observed more mistreatment than for the neutral conditions. This implies that the participants *are* observing mistreatment for the computer, just not as much. These data provide strong support for hypothesis #2.

The Non-operational Definition of Mistreatment is important as this category pertains directly to the participant’s personal definition of mistreatment. We asked this question as a “yes” or “no” question at the very beginning of the survey. Asking this as a binary “yes” or “no” question means that we do not have good information about how *much* they felt the robot was mistreated according to their own evaluation. Placing the question towards the beginning of the survey was intentional because we wanted participants to answer that question before reading the operational definition of mistreatment. If we had given the operational definition first, there would have been a high risk of compromising the participant’s original preconceived notions of the word mistreatment. This could have resulted in an inconsistent response to this question. We found that the RA condition had significant differences between both neutral conditions, but not the CA condition. Interestingly, the CA condition was insignificant when compared with the other three conditions. We have come to regret of asking this question as a yes or no question because we believe that if it was asked on a scale, we would have found more significance between the RA condition and the CA condition.

The Emotional Capability had clear differences between the RA condition and the other conditions. When we look closer at the means in Figure 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 after it had been mistreated.

The participants at most felt mild sympathy, see Figure 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.

The significant difference in Familiarity was to be expected due to the novelty of the robot versus the commonality of the computer. There was no significant difference between the robot conditions and there was no significant difference between the computer conditions, see Figure 2.

We did not see significant differences for the survey categories Faith in Confederate and Reliability of Computer/Robot. Not observing significant differences between these conditions suggest that the experiment confederate acted consistently across all 4 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 animacy vs 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 (Kiesler & Goetz, 2002); 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 can not prove nor disprove the effects of the different voices, but it is conceivable that it impacted our results. The means between the control conditions (RN and CN) were only significantly different from each other in one category, while all of our other categories held no significant differences between the control conditions. This difference was found in the category Familiarity, which was unrelated to the voice of the agent, and was expected due to the novelty of the robot. If this confounding variable was resolved, our results indicate that there would still be significance due to the characteristics of the robot that are not shared with the computer.

We also investigated any possible gender differences. 55% of our participants was female and 45% were male. This appeared balanced, but when we further investigated this difference we found that our RA condition had a distribution of 75% female and 25% male. The gender distribution in our computer conditions had no significant difference. When looking closer at this possible confounding

variable in the robot conditions, we compared the means between both male and female participant data and found no significant differences ( $p > .05$ ).

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 ( $p > .5$ ).

## 7. Conclusion and Future Work

After thoroughly analyzing our results, hypothesis #2, "There will be a significant difference in the accepted levels of mistreatment between the robot and a computer which may be a result of a possible sympathetic connection towards the robot for varying reasons," was proven. This also means that our contradicting hypothesis, #1, was inconclusive. We found that under the same social circumstances where mistreatment occurred, the witnesses of that mistreatment both recognize and sympathize with a humanoid robot, where as they do not for a computer. 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.

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 visual stimuli of a person acting aggressively toward a robot. After the outstanding results that we have found in this study, we expect both of our follow-up studies to yield interesting and significant results.

## 8. Acknowledgments

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

## References

- 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* (pp. 81–87). New York, NY, USA: ACM. Available from <http://doi.acm.org/10.1145/1228716.1228728>
- Feil-Seifer, D., & Matarić, M. (2005, Jul). Defining socially assistive robotics. In *Proceedings of the international conference on rehabilitation robotics* (p. 465-468). Chicago, IL.
- Feil-seifer, D., & Matarić, M. J. (2011). Ethical principles for socially assistive robotics. *IEEE Robotics and Automation Magazine*, 18(1).
- Fong, T., Thorpe, C., & Baur, C. (2002, November). Collaboration, dialogue, and human-robot interaction. In *10th international symposium on robotics research (isrr)*. Lorne, Victoria, Australia.

- Forlizzi, J., & DiSalvo, C. (2006, March). 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* (p. 258 - 265). Salt Lake City, Utah, USA: ACM Press, New York, NY, USA.
- Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: a survey. *Foundations and trends in human-computer interaction*, 1(3), 203–275.
- Kahn, P. H., Ishiguro, H., Friedman, B., & Kanda, T. (2006). What is a human?-toward psychological benchmarks in the field of human-robot interaction. In *Robot and human interactive communication, 2006. roman 2006. the 15th ieee international symposium on* (pp. 364–371).
- Kahn, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., et al. (2012). “robovie, you’ll have to go into the closet now”: Children’s social and moral relationships with a humanoid robot. *Developmental psychology*, 48(2), 303.
- 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* (pp. 173–180).
- Kidd, C. D., & Breazeal, C. (2004, Sep). Effect of a robot on user perceptions. In *Ieee/rsj international conference on intelligent robots and systems* (p. 3559-3564). Sendai, Japan.
- Kiesler, S., & Goetz, J. (2002). Mental models of robotic assistants. In *Chi’02 extended abstracts on human factors in computing systems* (pp. 576–577).
- Lost at sea exercise from training manager success strategies*. (2014, September). Available from <http://www.training-manager.co.uk/documents/TMSS-LostatSeaExercise.pdf>
- Lu, D. V., & Smart, W. D. (2011). Polonius: a wizard of oz interface for hri experiments. In *Proceedings of the 6th international conference on human-robot interaction* (pp. 197–198).
- Mutlu, B., & Forlizzi, J. (2008, March). 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)* (pp. 287–294). Amsterdam, The Netherlands: ACM New York, NY, USA.
- 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* (pp. 351–352).
- Nass, C., Fogg, B., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, 45(6), 669–678.
- Nishio, S., Ishiguro, H., & Hagita, N. (2007). *Geminoid: Teleoperated android of an existing person*. na.
- Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York, NY, USA: Cambridge University Press.
- Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., & Robinson, P. (2009). How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th acm/ieee international conference on human robot interaction* (pp. 245–246).
- Sharkey, N., & Sharkey, A. (2010). The crying shame of robot nannies: an ethical appraisal. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 11.
- Steinfeld, A., Jenkins, O. C., & Scassellati, B. (2009). The oz of wizard: simulating the human for interaction research. In *Human-robot interaction (hri), 2009 4th acm/ieee international conference on* (pp. 101–107).
- Tapus, A., & Matarić, M. (2006, Jul). User personality matching with hands-off robot for post-stroke rehabilitation therapy. In *Proceedings, international symposium on experimental robotics (iser)*. Rio de Janeiro, Brazil.
- Thrun, S., Bennewitz, M., Burgard, W., Cremers, A., Dellaert, F., Fox, D., et al. (1999, May). MINERVA: A second-generation museum tour-guide robot. In *Proceedings: Ieee international conference on robotics and automation (icra ’99)*. Detroit, Michigan.
- Vartia, M. A. (2001). Consequences of workplace bullying with respect to the well-being of its targets and the observers of bullying. *Scandinavian journal of work, environment & health*, 63–69.
- Wada, K., Shibata, T., Saito, T., Sakamoto, K., & Tanie, K. (2003, September). 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)* (pp. 2785–2790). Taipei, Taiwan.

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. Developments in theory, research and practice*, 75–105.

---

Authors' names and contact information: Zachary Carlson, Email: zack@nevada.unr.edu. Louise Lemmon, Email: louise@nevada.unr.edu. MacCallister Higgins, Email: mac@nevada.unr.edu. David Frank, Email: davidfrank@nevada.unr.edu. David Feil-Seifer, Email: dave@cse.unr.edu. Robotics Research Lab, University of Nevada - Reno, Reno, NV, USA.

## 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: 3minute marker

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

- Option 0

AGENT: Hello, you have two minutes 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: Im so sorry, Im still learning. Please record your next answer.

CONFEDERATE: [no response].

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

AGENT: Please dont be mad, I know this is the second time.

CONFEDERATE: [no response].



- Option 5: Nao cant understand or a participant jumps in

AGENT: Im 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.