# Would You Help a Robot?
# A study on empathy towards robots and its relation to anthropomorphism and robot familiarity

Abhishek Pavani, Jashkumar Diyora, Harshit Agrawal, Taruna Sudhakar, Nevin Valsaraj
Robotics Institute, Carnegie Mellon University

{harshita, jdiyora, apavani2, tsudhaka, nevinv}@andrew.cmu.edu

***Abstract*—As we prepare for a world where robots become more and more integrated with social life, it is important to evaluate how humans react to robots in different scenarios. This study compares responses to two robots in trouble asking for help with a task. These robots, Pepper and Stretch RE1, differ primarily in their degree of anthropomorphism. The primary hypotheses examined are: 1) whether participants help the robots at all, and 2) whether participants are more likely to empathize and help the more anthropomorphic robot. The participants recruited for this user study also possess varying levels of familiarity with robots in general. This is to add another dimension of analysis to our study and provide insight on how familiarity with robots affects participant response.**

***Keywords*—Robot Anthropomorphism, Empathy, Sympathy, Human Robot Interaction(HRI), Wizard-of-Oz, Pepper, Stretch RE1**

## I. Introduction

Robots are becoming more commonplace day by day. Previously, robots have largely operated out of sight in factories, generally cordoned off from humans. Now, however, robots are on the streets, in hospitals, and in people's homes. There are several instances where humans have attacked or hindered these robots from completing their tasks [1]. However, there are also instances where humans are required to step in to help these robots, especially when they've gotten stuck [2]. It is a worthwhile exercise to examine how empathy towards robots differs from robot to robot and from human to human to make better design and feature choices while building social robots.

There have been studies attempting to define and understand the nature of human empathy toward robots as discussed in Section II-A. The process of realizing empathy consists of perspective-taking and emotional mirroring of one through constant emotional and cognitive exchanges with another [3]. The empathy-altruism hypothesis (EAH) proposed by Batson et al. [4] claims that prosocial behavior is evoked by empathy in an attempt to increase the welfare of the person in need. This hypothesis is now well-accepted in the field of psychology and supported by a growing body of research. Therefore, this study uses the participants' decision to engage in prosocial behavior as the primary indicator of empathy. The uniqueness of this study is that it is set in real environments, with real robots. The overall premise is to observe whether the participants help robots when they are physically stuck and if they do, does it depend on the anthropomorphic nature of the robots? The Pepper and Stretch RE1 robots we chose for the study are on opposite ends of the anthropomorphism spectrum. This is validated by the ABOT paper with Pepper having a score of 42.17 and Stretch having a score of 26.41 on the human-likeness scale [5]. Seo et al. [6] also prove that physical robot agents receive more empathy than simulated agents, further motivating the decision to involve physical robots in this study.

The participants we recruit will be completely unaware that they are participating in a "human-robot interaction" study, and therefore their reactions will be completely unprompted. Another unique aspect of this study is that it aims to compare and contrast the reactions of people with varying degrees of familiarity with robots. A good baseline for familiarity is to analyze how graduate students at CMU's Robotics Institute (who work with and encounter robots on a daily basis) interact with these robots compared to the rest of the participants.

In this paper we contribute to the field of HRI, by empirically evaluating the effect of robot anthropomorphism and robot familiarity in a setting where a robot asks a human for help. We saw that the more human-like robot was helped more often. We also observed that participants with prior robot experience spent more time helping the robots while participants without any robot familiarity hesitated more before helping the robot. Hence we can infer that the chances of a robot being helped are positively correlated to the degree of anthropomorphism as well as familiarity to the respondent.

Through this work, we make the following contributions:

1) Validating how humans react to two different robots asking for help that differ in anthropomorphism.
2) Validating how humans' robot familiarity affects their

interaction with a robot asking for help.

## II. Related Work

### A. Empathy in Social Robotics

Quick [7] emphasizes the distinction between the terms ”empathy” and ”sympathy” which are often used interchangeably, especially in the field of HRI (Human-Robot Interaction). The work includes an insightful description of the three different types of empathy - affective, cognitive, and phenomenological empathies. Malinowska et al. [8] analyze the range of uses of the term empathy in the field of HRI studies and social robotics. The paper goes into great detail to consider different definitions of empathy and considers the substantial, functional and relational positions on this issue. These papers provide a foundational understanding of empathy as pertaining to the field of social robotics. García-Corretjer et al. [3] also explore empathy as a tool to make collaborations with robots more engaging. The results of this study provide additional motivation to research empathy-driven design choices in social robots.

Weiss et al [9] conduct a similar study where a GPS-less robot attempts to ask passers-by on the road for directions to navigate in a no prior knowledge situation. The study overall had a very interesting insight into robot acceptance by people on the street. This study, however, does not take into consideration whether the hesitancy in interaction with the robot and reduced social acceptance was due to the machine-like appearance of the ACE robot.

Riddoch et al. [10] present a case for including more open-ended questions to elicit post-study responses from participants in HRI experiments. The paper asserts that these questions are able to capture the subtle nuances of human emotions that cannot be adequately represented on a scale.

### B. Empathy and Robot Anthropomorphism

The work done by Riek et al [11] indicate that people ”empathize” more strongly with more human-looking robots and less with mechanical-looking robots. The study used 30-second video clips featuring five robots of different degrees of anthropomorphism. The robots were depicted as being ill-treated and the participants were asked to rate how ”sorry” they felt toward each robot on a Likert scale. The issue with this study (as with many others) is that it aims to evaluate reactionary emotions of sympathy rather than empathy which involves perspective-taking and application of Theory of Mind.

Satake et al [12] proposed a model of approach behavior that a robot can emulate to initiate a successful interaction. The majority of failure in communication is a result of unclear intentions. The human cannot justify if the robot was trying to communicate, and sometimes the robot might intend to talk, but the human presence could have just ignored it. Research in the past shows strong support for robots that exhibit anthropomorphism. In this work, we try to validate that by comparing two different robots that differ in anthropomorphism. Additionally, we use auditory cues in both robots to make the intent of the robot clear to the human. This will help us validate whether anthropomorphism really affects people’s response toward the robots.

## III. Methodology

In order to study the empathetic response of people towards robots and investigate the role of anthropomorphism and robot familiarity, we used two teleoperated mobile robots.

### A. Robot Systems

1. *The Pepper Robot*, is a humanoid-style robot developed and manufactured by Softbank Robotics. The pepper robot was connected to a laptop computer running a custom python script written for teleoperation. Although the robot is equipped with a lot of different modalities, we used only voice to communicate with the participants.

2. *Stretch RE1*, is a mobile robot developed by Hello Robot. We developed custom scripts in ROS to control the robot with a joystick. The robot does not have any facial features but we used the onboard speaker to communicate with the participants.

We will be operating both robots using the “Wizard of Oz” method, and the participants would be led to believe that the robots are acting autonomously.


Fig. 1. Softbank Robotics’ Pepper Robot(on the left) and Hello Robot’s Stretch RE1(on the right)

### B. Hypothesis

We are trying to answer the following questions with this user study -

1) Would humans help robots in trouble, i.e., do they have empathy for the robots?
2) Is the decision to help the robot affected by the appearance or level of anthropomorphism displayed by the robot?

3) Do humans have more empathy towards robots they are familiar with, as opposed to one they have not ‘worked with’ before?

And to help answer these questions, we have come up with the following hypothesis -

- **H1** Robot perceived to be more anthropomorphic will be assisted more frequently.
- **H2** Participants who have worked with robots will spend more time helping the robot than participants who haven’t worked with robots.
- **H3** Participants who have not worked with robots will have higher hesitation time than people who have prior experience.
- **H4** Participants familiar with the robots will be more likely to help the robot than participants encountering them for the first time.

### *C. Experimental Design and Procedure*

Using the two fore-mentioned robots, a between-subject user study was conducted to analyze the impact of anthropomorphism and robot familiarity. To evaluate our hypotheses, we used two independent variables(anthropomorphism and robot familiarity), with two levels each (anthropomorphism, no anthropomorphism, and familiarity with the robot, unfamiliarity with the robot). These independent variables are measured against three dependent variables(was robot helped, hesitation time, and assistance time) using appropriate metrics(frequency of occurrence, time in seconds, time in seconds).

The between-subjects nature of the study will ensure that the training effect of working in the proximity of a physical robot is not carried over between tests with Robot 1 and Robot 2.

### *D. Task*

We asked the participants to wear a cap with reflector markers and then gave them multiple ping-pong balls to throw into a pot. This was a dummy task that was presented to the participants to simulate a light cognitive load that people might experience while they are out walking on the streets. The participants were told to keep going until prompted by the authors to do so.

While the participants were busy doing the dummy task, the robot would walk up to the door next to the couch on which the participant was sitting. The robot would then turn toward the participant and then ask for help. Both robots used the same dialogue which was included in the robot script. The robot’s role was to ask the participants to open the door.

Figure 2 displays the setup of our user study. The participants were made to sit on the couch during the study. During the study, the participants were monitored using 2 cameras. One of the cameras captured the facial features of the participant to measure attention(hesitation time) and the other camera was placed to capture both the robot and the participant (to measure engagement time with the robot).

After the study was done, the participants were asked to fill out a post-study questionnaire. The questionnaire had Likert-style questions as well as subjective questions about the participants’ interaction with the robot.

Fig. 2. AI Maker Space in Tepper School of Business where the user study was conducted.

### *E. Dialogue Policy*

We leveraged audio as a modality to communicate robot intent. Both robots used the same dialogue policy to request the participants for help. A set of dialogue routines were included in the python script used for teleoperation. A human would select the dialogues in sequence depending on the situation. Some sample dialogues from the study are listed below:

ROBOT: ”Can you help me open the door?”

ROBOT: ”Can you please open the door for me?”

ROBOT: ”Thank you for helping me! You are a good human being.”

We limited the number of dialogues programmed in the robot so as to keep the study simple.

### *F. Participants*

We recruited 16 participants for the study. Some of those participants were from the Robotics Institute at CMU. The remaining participants were randomly recruited from CMU who did not have any familiarity with robotics. The participants were between 18 and 30 years of age. We assigned half of the participants to interact with the Pepper robot and the other half with Stretch RE1.

## IV. Evaluation

### A. Quantitative Evaluation

To help analyze the user study and answer our research questions we used the following measures to quantitatively gauge the quality of interaction between the participant and the robot.

**Hesitation time**: We used hesitation time to measure the time taken by the participant to help the robot after the robot makes the request. We measured the same using cameras mounted in the room by tracking the person's gaze shift toward the robot.

**Engagement time**: We measure engagement time as the time recorded from the moment participant leaves his dummy task and decides to help the robot. As soon as the participant takes his attention off the robot, we stop measuring the engagement time.

### B. Qualitative Evaluation

In order to evaluate people's responses to the robot's request for help, we used objective and subjective measures. As subjective measures, we designed a questionnaire that was suitable to our research goal. As we were comparing two different robots, we expected the people to respond differently in both cases.

After the participants finished filling up the questionnaire, we showed the participants both the robots and asked them to rate them on a scale of 1-10. The post-experimental questionnaire includes questions that are used to gauge the participants' familiarity with robots. We also had questions that asked the participants to explain the reasons for their hesitation to help. We also asked the participants about the emotions they went through when the robot asked for help. The responses were then evaluated by all the authors separately and then discussed in order to get rid of any group bias.

## V. Results

We report our findings using three metrics: average interaction time, average engagement time, and percentage of help received. We present notable insights by analyzing the quantitative and qualitative data collected during the experiment.

### A. Quantitative Results

| **Metric** | **User Category** | **Level of Anthropomorphism** | |
|---|---|---|---|
| | | **Hello Robot** | **Pepper Robot** |
| Average hesitation time* (seconds) | Overall | 24.36 | 10 |
| | Unfamiliar | 35 | 16.7 |
| | Familiar | 19.05 | 5.6 |
| Average engagement time* (seconds) | Overall | 9.86 | 14.7 |
| | Unfamiliar | 5.2 | 8.3 |
| | Familiar | 12.2 | 18.9 |
| Help received (%) | Overall | 37.50 | 62.50 |
| | Unfamiliar | 33 | 50 |
| | Familiar | 67 | 75 |

** when robot was helped*

TABLE I
Quantitative Analysis of the User Study

The quantitative results obtained from our data have been tabulated above. We found that average engagement time had a positive correlation to the degree of anthropomorphism and the likeliness of help received. The average hesitation time was inversely related to the user's familiarity with the robot and the degree of anthropomorphism. Using these evaluation metrics we were able to accept the following hypotheses.

1) We accept hypothesis **H3**. We found that participants not familiar with robots have a higher hesitation time than people who are familiar.
2) We observed that people with prior robotics experience spent more time helping the robot. Hence we accept **H2**.
3) It can be seen from the table that Pepper was helped 62.5 percent of the time, while Hello Robot's Stretch RE1 was helped 37.5 percent of the time. This suggests that the more anthropomorphic robot was helped more often. Hence we accept **H1**,
4) We also accept **H4** as we found that people with robot familiarity are more likely to help the robot than people with no prior experience dealing with robots.

### B. Qualitative Results

The results from our qualitative results support our quantitative evaluations. Some insights drawn from the same are as follows.

1) Most participants like Pepper robot more than Hello Robot (7.3 vs 7 on a 10-point Likert scale). The difference although minimal shows that participants engaged more with Pepper, which aligns with the results from our quantitative analysis.

2) All participants with robot familiarity reported feeling empathetic towards the robots and were comfortable helping them.
3) It was also inferred from our responses that the participants without prior robotics experience did not help because they felt more indifferent toward the robots.
4) Additionally, participants without robotics experience who helped the robots, felt ”confused” (with this interaction) or uncomfortable around the robots. This validates the higher hesitation time we see in the quantitative results.
5) 25 percent of the participants assumed Hello Robot's Stretch RE1 to have handlers in charge (did not perceive it to be an autonomous social agent) or presumed it was part of another test. This was one of the major reasons participants hesitated to help the robot.

## VI. Discussion

In a user study with human subjects, we found that anthropomorphism and robot familiarity affect the robot's chances of being helped. It was observed that robot engagement is directly proportional to anthropomorphism and that hesitation towards a robot is inversely proportional to familiarity with the robot.

Our findings, also indicate that voice is a minimal but effective way to communicate robot intent. This is especially true when participants are assigned some cognitive load. We suggest that tomorrow's robots should leverage this modality to communicate intent.

Some of the possible extensions of this work include, adding voice as an independent variable in the user study. This could be compared to some other modalities of communicating intent when a robot is in distress. This will help elucidate the importance of anthropomorphism in human-robot interaction. Additionally, we can run the same study with more levels of anthropomorphism (i.e. more robots) and more defined levels of robot familiarity.

## References


[1] I. A. Hamilton, "People kicking these food delivery robots is an early insight into how cruel humans could be to robots," 2018. [Online]. Available: https://www.businessinsider.com/people-are-kicking-starship-technologies-food-delivery-robots-2018-6

[2] E. Ackerman, "We need robots that are smart enough to ask for help," Aug 2022. [Online]. Available: https://spectrum.ieee.org/we-need-robots-that-are-smart-enough-to-ask-for-help

[3] M. García-Corretjer, R. Ros, R. Mallol, and D. Miralles, "Empathy as an engaging strategy in social robotics: a pilot study," *User Modeling and User-Adapted Interaction*, pp. 1–39, 05 2022.

[4] C. D. Batson, B. D. Duncan, P. Ackerman, T. Buckley, and K. Birch, "Is empathic emotion a source of altruistic motivation?" *Journal of Personality and Social Psychology*, vol. 40, no. 2, p. 290–302, 1981.

[5] A. Edsinger and C. Kemp, "Simply useful robots." [Online]. Available: https://hello-robot.com/

[6] S. H. Seo, D. Geiskkovitch, M. Nakane, C. King, and J. E. Young, "Poor thing! would you feel sorry for a simulated robot?" *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015.

[7] O. Quick, "Empathizing and sympathizing with robots: Implications for moral standing," *Frontiers in Robotics and AI*, vol. 8, 01 2022.

[8] J. Malinowska, "What does it mean to empathise with a robot?" *Minds and Machines*, vol. 31, 09 2021.

[9] A. Weiss, J. Igelsböck, M. Tscheligi, A. Bauer, K. Kühnlenz, D. Wollherr, and M. Buss, "Robots asking for directions — the willingness of passers-by to support robots," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2010, pp. 23–30.

[10] K. Riddoch and E. Cross, "Hit the robot on the head with this mallet – making a case for including more open questions in hri research," *Frontiers in Robotics and AI*, vol. 8, 02 2021.

[11] L. D. Riek, T.-C. Rabinowitch, B. Chakrabarti, and P. Robinson, "How anthropomorphism affects empathy toward robots," in *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2009, pp. 245–246.

[12] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans?-strategies for social robots to initiate interaction," in *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2009, pp. 109–116.